

# UNIVERSITÉ PARIS DESCARTES

École Doctorale GC2ID

---

THÈSE

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ PARIS DESCARTES**

Sciences de la Vie et de la Santé

Discipline : Génétique Humaine

**RECHERCHE DE NOUVEAUX FACTEURS GÉNÉTIQUES DE  
SUSCEPTIBILITÉ À LA SPONDYLOARTHRITE GRÂCE À UNE APPROCHE  
ASSOCIANT ÉTUDES FAMILIALES ET GÉNOMIQUE FONCTIONNELLE**

Présentée et soutenue publiquement par

**Félicie COSTANTINO**

Le 7 novembre 2014

## JURY

Mr le Pr. Jean ROUDIER	Président du jury
Mr le Pr. Arnaud CONSTANTIN	Rapporteur
Mme le Pr. Elisabeth PETIT-TEXEIRA	Rapporteur
Mr le Pr. Henri-Jean GARCHON	Examineur
Mr le Pr. Maxime BREBAN	Directeur de thèse

## RÉSUMÉ

---

La spondyloarthrite (SpA) est un rhumatisme inflammatoire chronique fréquent et invalidant. Plus d'une vingtaine de locus de susceptibilité à la maladie ont été identifiés à ce jour, dont HLA-B27 situé dans le complexe majeur d'histocompatibilité (CMH). L'objectif de ce travail était d'identifier de nouveaux facteurs génétiques de susceptibilité à la SpA grâce à une double approche d'études familiales et de génomique fonctionnelle.

Dans la première partie, nous avons génotypé des familles multiplex de SpA. L'analyse de liaison non paramétrique a révélé la présence, en plus du CMH, d'un nouveau locus significativement lié à la SpA en 13q13. L'étude de ce locus nous a permis de restreindre la région d'intérêt à un intervalle de 1,3 Mb, dont le séquençage est en cours. Par ailleurs, l'étude d'association intra-familiale a identifié un SNP intronique de *MAPK14* significativement associé à la SpA. Enfin, nous avons montré que l'un des SNPs du gène *IL23R* connu pour être associé à la spondylarthrite ankylosante était en fait associé à la présence d'une sacro-iliite radiologique dans la SpA.

Parallèlement aux études familiales, nous avons comparé le transcriptome de cellules dendritiques de neuf patients atteints de SpA à celui de dix témoins sains. Nous avons ainsi identifié 81 gènes différemment exprimés. Nous avons aussi montré que l'expression génique d'*ERAP1* (et à un moindre degré son expression protéique et son niveau d'activité enzymatique) étaient sous le contrôle de polymorphismes de ce gène associés à la SpA.

Mots clés : spondyloarthrite, génétique, étude de liaison, association intra-familiale, transcriptomique, cellules dendritiques, eQTL

## ABSTRACT

---

Spondyloarthritis (SpA) is a frequent and disabling chronic rheumatic disease. To date, more than 20 susceptibility loci have been identified, including HLA-B27 in the major histocompatibility complex (MHC). Most of the disease heritability remains to be elucidated. The aim of the study was to identify new genetic factors of susceptibility to SpA using an approach combining genetics and functional genomics.

In the first part of this work, we genotyped SpA multiplex families with microarrays of 250,000 SNPs. Non parametric linkage analysis revealed a new locus significantly linked to SpA outside the MHC, in 13q13. Further studies on this locus allowed us to map the disease interval to a 1.3 Mb region, which will be soon sequenced. Moreover, family-based association study identified a significant association between one intronic SNP in *MAPK14* and SpA. We also showed that one of the known ankylosing spondylitis-associated SNP in *IL23R* was indeed associated with sacroiliitis in SpA.

We have also compared dendritic cells gene expression between nine SpA patients and ten controls and identified 81 genes differentially expressed. Moreover, we showed that *ERAP1* gene expression (and at a less extent protein expression and enzymatic activity) is under the control of several polymorphisms in the gene which has previously been associated with SpA.

Keywords: spondyloarthritis, genetics, linkage analysis, family-based association, transcriptomics, dendritic cells, eQTL

## REMERCIEMENTS

---

Au terme de ces quatre années, il est temps de remercier l'ensemble des personnes qui ont participé de près ou de loin à cette « aventure », en espérant n'oublier personne.

Mes premiers remerciements vont à mon directeur de thèse, Maxime BREBAN, pour m'avoir confié ce sujet passionnant et donné les moyens de réaliser ce travail. Grâce à votre enthousiasme sans faille et à la confiance que vous m'avez accordée, j'ai pu traverser sans encombre les moments les plus difficiles. J'espère avoir été à la hauteur de vos attentes et continuer encore longtemps à travailler à vos côtés.

Je remercie Henri-Jean GARCHON pour avoir co-dirigé ce travail et avoir su faire preuve de patience et de pédagogie pour m'expliquer les statistiques et la génétique. Votre rigueur et vos compétences ont été primordiales à l'avancée de cette thèse.

Je tiens à remercier tous les membres de mon jury de thèse. Merci au professeur Jean ROUDIER de me faire l'honneur d'en être le président. Je remercie les professeurs Elisabeth PETIT-TEXEIRA et Arnaud CONSTANTIN d'avoir pris le temps de rapporter ce travail, je les remercie également pour leurs remarques constructives sur le manuscrit.

Ma reconnaissance va aussi à Gilles CHIOCCHIA, co-directeur du laboratoire, pour m'avoir accueillie au sein de l'équipe et pour son regard toujours positif et constructif.

Ce travail n'aurait pas pu voir le jour si d'autres avant moi n'avaient pas ouvert le chemin : je remercie Elena et Amir qui, grâce à la qualité de leurs travaux, ont permis de construire ce projet sur des bases solides. Merci pour votre aide précieuse, notamment lors de mes débuts sur ce projet.

Rien n'aurait été possible sans l'ensemble des patients et de leurs familles qui ont accepté de participer aux études génétiques et sans le travail méticuleux d'inclusion de Roula et d'Ariane.

Un grand merci à ceux avec qui j'ai travaillé au quotidien. Un merci tout particulier à Alice : travailler avec toi a toujours été un plaisir grâce à ton enthousiasme et ta bonne humeur jamais mise à mal ! Nelly, merci de m'avoir initiée (bien que brièvement) au L2 et surtout pour le café du matin et le RER du soir ! Merci à Brigitte, Emmanuel, Tifenn, mais également aux étudiants de passage (Karim, Zeyna, Camille).

Une pensée pour toute la plateforme de génomique, et tout particulièrement pour Sébastien, pour l'aide qu'ils m'ont fourni tout au long du projet.

Je n'oublie bien sûr pas tous les pensionnaires passés ou présents du laboratoire, que ce soit à Cochin ou à Montigny : mes colocataires de couloir (Emna, Sandrine), puis de bureaux (Amir, Luiza, Olivier, Tifenn, Valérie) et tous les autres : Abdel, Adèle, Aurélie, Aurore, Catherine, Cindy, Claudine, Greg, Ingrid, Isabelle, Marie-Agnès, Nathalie, Ourdia, Roméro, Sandrine, Simon, Sylvie, sans oublier la « Giraud » team (Clotilde, Mathieu, Nada), la « Allanore » team (Barbara, Muriel, Eugénie, Maxime) et bien sûr la plateforme d'imagerie du petit animal (Didier, Gilles, Carmen, Frank). Et tous les autres que j'oublie...

Autres temps, autres lieux : merci à tous mes anciens collègues du CHU de Nancy et à l'ensemble du service de rhumatologie d'Ambroise Paré (et tout particulièrement à Gilles et Maria-Antonietta) pour cette première année passée à leurs côtés, en espérant qu'il y en ait de nombreuses autres !

Enfin, des remerciements tout particuliers à ma famille et mes amis pour le soutien et l'aide qu'ils ont pu m'apporter au quotidien. Merci notamment aux relecteurs attentifs du manuscrit (ils sauront se reconnaître...).

Merci à tous !

# TABLE DES MATIÈRES

RÉSUMÉ .....	1
REMERCIEMENTS.....	2
TABLE DES MATIÈRES.....	4
LISTE DES FIGURES ET DES TABLEAUX .....	8
LISTE DES ABRÉVIATIONS .....	9
PREMIÈRE PARTIE : INTRODUCTION GÉNÉRALE.....	11
Chapitre 1. La spondyloarthrite : aspects épidémiologiques, cliniques et thérapeutiques.....	12
I.1. Épidémiologie .....	12
I.1.1 Prévalence .....	12
I.1.2 Sexe ratio.....	12
I.1.3 Âge de survenue.....	13
I.2. Diagnostic et traitement.....	14
I.2.1 Manifestations cliniques .....	14
I.2.1.1 Manifestations articulaires.....	14
I.2.1.2 Manifestations extra-articulaires .....	16
I.2.2 Examens complémentaires .....	18
I.2.2.1 Examens biologiques .....	18
I.2.2.2 Imagerie.....	18
I.2.3 Traitement.....	20
I.2.3.1 Traitements médicamenteux .....	21
I.2.3.2 Traitements physiques .....	22
I.2.3.3 Traitements chirurgicaux.....	22
I.3. Concept unifié de spondyloarthrite .....	23
I.3.1 Les sous-types de spondyloarthrite.....	23
I.3.1.1 Spondylarthrite ankylosante .....	23
I.3.1.2 Rhumatisme psoriasique .....	23
I.3.1.3 Arthrites réactionnelles.....	24
I.3.1.4 Arthrites associées aux maladies inflammatoires chroniques de l'intestin .....	24
I.3.1.5 Spondylarthrites indifférenciées .....	24
I.3.2 La spondyloarthrite, un regroupement justifié.....	24
I.3.3 Critères de classification de la spondyloarthrite .....	25
I.3.3.1 Critères d'Amor .....	25
I.3.3.2 Critères de l'European Spondyloarthropathy Study Group .....	25
I.3.3.3 Critères de l'Assessment of SpondyloArthritis international Society .....	27
Chapitre 2. Méthodologie de la génétique des maladies complexes.....	29
II.1. Caractérisation de la variabilité génétique humaine .....	29
II.1.1 Microsatellites .....	29
II.1.2 Polymorphismes mono-nucléotidiques.....	30

II.1.3 Variations du nombre de copies.....	30
II.2. Analyses de liaison génétique .....	31
II.2.1 Analyses de liaison paramétriques.....	31
II.2.1.1 Calcul du LOD score .....	32
II.2.1.2 Prise en compte de l'hétérogénéité de locus : LOD score d'hétérogénéité .....	33
II.2.1.3 Interprétation des résultats .....	34
II.2.1.4 Analyses paramétriques et maladies complexes .....	34
II.2.2 Analyses de liaison non paramétriques.....	35
II.2.2.1 Méthode des paires de germains atteints .....	35
II.2.2.2 Extension à d'autres types de structures familiales.....	36
II.2.2.3 Interprétation des résultats .....	37
II.3. Analyses d'association génétique.....	38
II.3.1 Déséquilibre de liaison .....	38
II.3.2 Tests d'association .....	40
II.3.2.1 Analyses cas-témoins .....	40
II.3.2.2 Études d'association intra-familiale .....	41
II.3.3 Stratégie générale .....	41
II.3.3.1 Approche gène-candidat .....	42
II.3.3.2 Approche génome entier .....	42
II.3.4 Contrôle-qualité des données .....	43
II.3.4.1 Évaluation de la qualité des échantillons d'ADN.....	43
II.3.4.2 Évaluation de la qualité des génotypes.....	43
II.3.4.3 Recherche d'un biais de stratification .....	44
II.3.5 Interprétation des résultats .....	45
II.3.5.1 Significativité et force de l'association .....	45
II.3.5.2 Correction pour tests multiples.....	46
II.4. Des polymorphismes génétiques à l'expression des gènes .....	46
II.4.1 Étude du transcriptome dans les maladies complexes .....	46
II.4.1.1 Définition .....	47
II.4.1.2 Techniques d'étude .....	47
II.4.1.3 Stratégie d'étude .....	47
II.4.2 Locus associé à l'expression d'un caractère quantitatif.....	48
II.4.2.1 Définition .....	48
II.4.2.2 Techniques d'étude .....	49
II.4.2.3 Stratégie d'analyse .....	50
Chapitre 3. La spondyloarthrite : une maladie génétique complexe .....	51
III.1. Épidémiologie génétique de la spondyloarthrite .....	51
III.1.1 Épidémiologie génétique de la spondylarthrite ankylosante.....	51
III.1.1.1 Agrégation familiale .....	51
III.1.1.2 Études de jumeaux .....	52
III.1.1.3 Études de ségrégation : modélisation de la transmission de la maladie .....	53
III.1.2 Épidémiologie génétique de la spondyloarthrite .....	53

III.1.2.1 Pertinence de l'étude de la spondyloarthrite dans son ensemble sur le plan génétique.....	54
III.1.2.2 Agrégation familiale dans la spondyloarthrite .....	54
III.1.3 Facteurs environnementaux .....	54
III.2. Facteurs de susceptibilité génétique identifiés dans la spondyloarthrite .....	55
III.2.1 Complexe majeur d'histocompatibilité .....	55
III.2.1.1 HLA-B27 .....	55
III.2.1.2 Autres gènes du complexe majeur d'histocompatibilité .....	56
III.2.1.3 Arguments pour l'implication d'autres facteurs génétiques en dehors du complexe majeur d'histocompatibilité .....	58
III.2.2 Locus de susceptibilité à la spondyloarthrite en dehors du complexe majeur d'histocompatibilité .....	58
III.2.2.1 Études de liaison.....	58
III.2.2.2 Études d'association.....	61
III.2.3 Contribution des études génétiques à la compréhension de la spondyloarthrite.....	64
III.2.3.1 Voies physiopathologiques impliquées.....	64
III.2.3.2 Spondylarthrite ankylosante, spondyloarthrite et autres maladies inflammatoires : un terrain génétique partagé .....	67
III.3. Étude du transcriptome dans la spondyloarthrite .....	69
III.3.1 Différences méthodologiques entre les études .....	70
III.3.1.1 Objectifs de l'étude .....	70
III.3.1.2 Choix du type cellulaire/tissulaire .....	70
III.3.1.3 Choix du comparateur .....	71
III.3.2 Contribution des études du transcriptome à la compréhension de la spondyloarthrite .....	72
OBJECTIFS DE LA THÈSE .....	74
DEUXIÈME PARTIE : TRAVAUX ORIGINAUX.....	76
Article 1 .....	77
Article 2 .....	103
Article 3 .....	119
Article 4.....	134
Article 5 .....	164
DISCUSSION & PERSPECTIVES.....	190
Études familiales : un outil pour l'identification de gènes de susceptibilité dans la spondyloarthrite.....	191
Positionnement du sujet et problématique générale .....	191
Résultats préliminaires et objectifs du projet .....	192
Analyse de liaison .....	193
Locus 13q13.....	193
Complexe majeur d'histocompatibilité .....	193
Comparaison avec les résultats de la littérature.....	194
Justification de la méthodologie et difficultés rencontrées.....	195
Perspectives.....	196

Analyse d'association .....	197
MAPK14 .....	198
Comparaison avec les résultats de la littérature.....	198
Perspectives.....	199
De la maladie au sous-phénotype : exemple du gène <i>IL23R</i> .....	199
Facteurs génétiques associés à l'atteinte radiologique .....	199
Facteurs génétiques associés à l'activité et à la sévérité fonctionnelle .....	200
Génomique fonctionnelle.....	200
Positionnement du sujet et problématique général .....	200
Étude du transcriptome des cellules dendritiques dans la spondyloarthrite .....	201
Cellules dendritiques et spondyloarthrite.....	201
Transcriptome des cellules dendritiques .....	202
Perspectives.....	203
Du polymorphisme à la fonction : exemple du gène <i>ERAP1</i> .....	204
Du polymorphisme à l'expression génique .....	204
Du gène à la protéine et à la fonction .....	204
Perspectives.....	205
QUATRIÈME PARTIE : CONCLUSIONS .....	206
CINQUIÈME PARTIE : BIBLIOGRAPHIE .....	208
ANNEXES .....	240
Annexe 1 .....	241
Annexe 2 .....	260
Annexe 3 .....	262



## LISTE DES FIGURES ET DES TABLEAUX

---

### FIGURES

<b>Figure 1</b> – Représentations des atteintes articulaires axiales dans la SpA.....	15
<b>Figure 2</b> – Atteintes des enthèses dans la SpA.....	16
<b>Figure 3</b> – Recommandations ASAS/EULAR de prise en charge de la SA .....	20
<b>Figure 4</b> – Différents types de marqueur génétique.....	31
<b>Figure 5</b> – Liaison génétique et fraction de recombinaison.....	32
<b>Figure 6</b> – Définition des principaux scores de liaison.....	33
<b>Figure 7</b> – Partage d'allèles entre germains - notion d'IBD et d'IBS.....	36
<b>Figure 8</b> – Calcul de la statistique NPL (Z score) .....	37
<b>Figure 9</b> – Associations directe et indirecte.....	39
<b>Figure 10</b> – Calcul des coefficients de DL $ D' $ et $r^2$ .....	39
<b>Figure 11</b> – Principes du TDT .....	41
<b>Figure 12</b> – Équilibre de Hardy-Weinberg.....	44
<b>Figure 13</b> – Biais de stratification .....	45
<b>Figure 14</b> – Calcul de l'odds ratio.....	45
<b>Figure 15</b> – Mécanismes des eQTLs .....	49
<b>Figure 16</b> – Les différentes étapes de l'épidémiologie génétique des maladies complexes.....	52
<b>Figure 17</b> – Implication des différentes aminopeptidases associées à la SA .....	65
<b>Figure 18</b> – Partage de locus de susceptibilité .....	69

### TABLEAUX

<b>Tableau 1</b> – Prévalence de la SpA à travers le monde.....	13
<b>Tableau 2</b> – Critères de New York modifiés de SA.....	19
<b>Tableau 3</b> – Critères de classification de SpA d'Amor .....	26
<b>Tableau 4</b> – Critères de classification de SpA de l'ESSG .....	26
<b>Tableau 5</b> – Critères de classification de SpA de l'ASAS .....	27
<b>Tableau 6</b> – Études de jumeaux réalisées dans la SA.....	53
<b>Tableau 7</b> – Principales hypothèses liant HLA-B27 à la SpA.....	57
<b>Tableau 8</b> – Résultats des analyses de liaison dans la SpA .....	60
<b>Tableau 9</b> – Principaux résultats des études d'association gène-candidat dans la SpA.....	63
<b>Tableau 10</b> – Résultats des études d'association génome-entier.....	64
<b>Tableau 11</b> – Résultats des études d'association comparant SA et SpA non radiographiques.....	68
<b>Tableau 12</b> – Partage des locus de susceptibilité entre la SA et d'autres pathologies .....	68
<b>Tableau 13</b> – Résultats des études du transcriptome dans la SpA .....	73

## LISTE DES ABRÉVIATIONS

---

- **ADN** : acide désoxyribonucléique
- **AINS** : anti-inflammatoire non stéroïdien
- **ANOVA** : analysis of variance (analyse de la variance)
- **Anti-TNF $\alpha$**  : anti-tumor necrosis factor  $\alpha$
- **ARé** : arthrite réactionnelle
- **ARN** : acide ribonucléique
- **ARNm** : acide ribonucléique messager
- **ASAS** : assessment of spondyloarthritis international society (société internationale d'évaluation de la spondyloarthrite)
- **ASDAS** : ankylosing spondylitis disease activity score (score d'activité de la spondylarthrite ankylosante)
- **BASDAI** : Bath ankylosing spondylitis disease activity index (indice d'activité de la spondylarthrite ankylosante de Bath)
- **BASFI** : Bath ankylosing spondylitis functional index (indice fonctionnel de la spondylarthrite ankylosante de Bath)
- **BASRI** : Bath ankylosing spondylitis radiological index (indice radiologique de la spondylarthrite ankylosante de Bath)
- **CMH** : complexe majeur d'histocompatibilité
- **CNV** : copy number variation (variation du nombre de copies)
- **CRP** : protéine C-réactive
- **DC** : dendritic cell (cellule dendritique)
- **DL** : déséquilibre de liaison
- **DZ** : dizygote
- **eQTL** : expression quantitative trait locus (locus associé à l'expression d'un trait quantitatif)
- **ERAP1** : endoplasmic reticulum aminopeptidase 1 (aminopeptidase 1 du réticulum endoplasmique)
- **ESSG** : European spondyloarthropathy study group (groupe d'étude européen des spondylarthropathies)
- **EULAR** : European league against rheumatism (ligue européenne contre les rhumatismes)
- **FDR** : false discovery rate (taux de fausse découverte)
- **FWER** : family wise error rate (taux d'erreur global)
- **GFECS** : groupe française d'étude génétique de la spondylarthrite
- **GWAS** : genome wide association study (étude d'association pangénomique)
- **HLA** : human leucocyte antigen (antigène des leucocytes humains)
- **HLOD** : heterogeneity logarithm of odds (LOD score d'hétérogénéité)
- **HW** : Hardy-Weinberg
- **IBD** : identical by descent (identique par descendance)
- **IBS** : identical by state (identique par état)
- **IC95%** : intervalle de confiance à 95%
- **IGAS** : international genetics of ankylosing spondylitis consortium (consortium international de la génétique de la spondylarthrite ankylosante)
- **IL23R** : récepteur de l'interleukine 23

- **IRM** : imagerie par résonance magnétique
- **Kb** : kilobase
- **KIR** : killer cell immunoglobulin-like receptors
- **LOD** : logarithm of odds
- **LPS** : lipopolysaccharide
- **LT** : lymphocyte T
- **LT-CD4+** : lymphocyte T CD4+
- **MAPK14** : mitogen-activated protein kinase 14
- **Mb** : mégabase
- **MD-DC** : monocyte-derived dendritic cell (cellule dendritique dérivée de monocyte)
- **MICI** : maladie inflammatoire chronique de l'intestin
- **MLR** : maximum likelihood ratio (rapport de vraisemblance maximal)
- **MZ** : monozygote
- **NASC** : North American spondyloarthritis consortium (consortium nord-américain de la spondyloarthrite)
- **NGS** : next generation sequencing (séquençage de nouvelle génération)
- **NK** : natural killer
- **NPL** : non parametric linkage (liaison non paramétrique)
- **OR** : odds ratio (rapport de cote)
- **PBMC** : peripheral blood mononuclear cell (cellule mononucléée du sang périphérique)
- **PCR** : polymerase chain reaction (réaction de polymérisation en chaîne)
- **pQTL** : protein quantitative trait locus (locus associé à un trait quantitatif protéique)
- **PR** : polyarthrite rhumatoïde
- **RhMICI** : rhumatisme associé aux maladies inflammatoires chroniques de l'intestin
- **RP** : rhumatisme psoriasique
- **SA** : spondylarthrite ankylosante
- **SNP** : single nucleotide polymorphism (polymorphisme mononucléotidique)
- **SpA** : spondyloarthrite
- **TDM** : tomodensitométrie
- **TDT** : transmission disequilibrium test (test de déséquilibre de transmission)
- **TNF $\alpha$**  : tumor necrosis factor  $\alpha$
- **UPR** : unfolded protein response
- **USpA** : spondylarthrite indifférenciée

## **PREMIÈRE PARTIE : INTRODUCTION GÉNÉRALE**

# CHAPITRE I. LA SPONDYLOARTHRITE : ASPECTS ÉPIDÉMIOLOGIQUES, CLINIQUES ET THÉRAPEUTIQUES

---

La spondyloarthrite (SpA) est un rhumatisme inflammatoire chronique fréquent. Ses manifestations cliniques variées et l'absence d'examen complémentaires de référence sont à l'origine de difficultés diagnostiques et thérapeutiques. En dépit d'une présentation clinique polymorphe, les différents sous-types initialement décrits sont désormais regroupés, notamment depuis l'élaboration de critères de classification communs.

## I.1. Épidémiologie

La SpA figure parmi les rhumatismes inflammatoires chroniques les plus fréquents en France et dans le monde. Sa grande hétérogénéité clinique rend toutefois difficile la comparaison entre elles des études épidémiologiques, que ce soit en terme de prévalence, de sexe ratio ou d'âge de survenue.

### I.1.1 Prévalence

De nombreuses études ont cherché à estimer la prévalence de la SpA (**Tableau 1**). On observe d'importantes variations en fonction de la population étudiée avec, d'une manière générale, une prévalence plus élevée dans les pays du Nord que dans les pays du Sud. Ces différences s'expliquent notamment par la répartition géographique de l'allèle HLA-B27, principal facteur de susceptibilité à la maladie<sup>1</sup>.

En France, la prévalence de la SpA a été évaluée à 0,3% dans une étude réalisée en 2001<sup>2</sup>. Cette prévalence est très proche de celle estimée la même année dans la polyarthrite rhumatoïde (PR) dans une étude utilisant la même méthodologie<sup>3</sup>. Plus récemment, notre équipe a estimé la prévalence de la SpA à 0,43% à partir de l'étude de la cohorte GAZEL dont les détails sont présentés dans l'**Annexe 1**<sup>4</sup>.

### I.1.2 Sexe ratio

Il est classiquement admis qu'il existe une prédominance masculine de la SpA, et plus particulièrement de la spondylarthrite ankylosante (SA), l'un des principaux sous-types de la maladie, avec un sexe ratio de l'ordre de 2:1 à 3:1<sup>5</sup>. Il est néanmoins intéressant de noter que ce déséquilibre de sexe a tendance à disparaître dans les études les plus récentes. Cette évolution suggère que les femmes atteintes de SpA ont eu tendance à être sous-diagnostiquées par le passé.

Cette hypothèse est étayée par les données d'une étude allemande portant sur plus de 1 300 patients où le pourcentage de femmes parmi les malades est passé de 10% en 1960 à 46% en 1990<sup>6</sup>. Ce sous-diagnostic pourrait s'expliquer par des différences de symptomatologie en fonction du sexe, avec notamment une plus grande sévérité radiologique (et donc un diagnostic plus aisé) chez les hommes<sup>5</sup>.

**Tableau 1 – Prévalence de la SpA à travers le monde ; d'après Stolwijk *et al.*<sup>7</sup>**

Étude	Pays	Année	Prévalence % (IC95%)
Boyer <i>et al.</i> <sup>8</sup>	Eskimos d'Alaska	1994	2,5
Alexeeva <i>et al.</i> <sup>9</sup>	Chukotka (minorité russe)	1994	2,5
Braun <i>et al.</i> <sup>10</sup>	Allemagne	1998	1,9
Farooqi <i>et al.</i> <sup>11</sup>	Pakistan	1998	0,10
Hukuda <i>et al.</i> <sup>12</sup>	Japon	2001	0,0095
Bruges-Armas <i>et al.</i> <sup>13</sup>	Espagne	2002	1,6 (0,8 – 2,7)
Minh Hoa <i>et al.</i> <sup>14</sup>	Viet Nam	2003	0,28
Saraux <i>et al.</i> <sup>2</sup>	France	2005	0,3 (0,17 – 0,46)
Trontzas <i>et al.</i> <sup>15</sup>	Grèce	2005	0,49 (0,38 – 0,60)
De Angelis <i>et al.</i> <sup>16</sup>	Italie	2007	1,06 (0,78 – 1,38)
Adomaviciute <i>et al.</i> <sup>17</sup>	Lituanie	2008	0,84 (0,53 – 1,21)
Davatchi <i>et al.</i> <sup>18</sup>	Iran	2008	0,23
Onen <i>et al.</i> <sup>19</sup>	Turquie	2008	1,05 (0,70 – 1,05)
Liao <i>et al.</i> <sup>20</sup>	Chine	2009	0,78 (0,62 – 0,95)
Haglund <i>et al.</i> <sup>21</sup>	Suède	2011	0,45 (0,44 – 0,47)
Reveille <i>et al.</i> <sup>22</sup>	États-Unis	2012	1,4 (1 – 1,9)
Costantino <i>et al.</i> <sup>4</sup>	France	2013	0,43 (0,26 – 0,70)

IC 95% : intervalle de confiance à 95%

### I.1.3 Âge de survenue

La SpA peut survenir à tout âge, mais elle affecte préférentiellement l'adulte jeune. L'âge moyen de survenue de la maladie est d'environ 25 ans<sup>6,23</sup>. Compte-tenu des difficultés diagnostiques, il existe souvent un délai entre les premiers symptômes et le diagnostic définitif<sup>24</sup>. Ce délai a toutefois tendance à diminuer comme en attestent plusieurs études récentes<sup>24,25</sup>.

La SpA est donc une maladie fréquente, touchant principalement le jeune adulte, sans prédominance de sexe. Le délai diagnostique, souvent très long, est en partie à mettre sur le compte de la grande variabilité des manifestations cliniques de la maladie.

## **I.2. Diagnostic et traitement**

Le diagnostic de SpA repose avant tout sur la mise en évidence d'un faisceau de signes cliniques évocateurs. Il peut en outre s'appuyer sur la réalisation d'examens complémentaires. La précocité du diagnostic est primordiale afin de proposer une prise en charge thérapeutique rapide permettant de limiter les conséquences fonctionnelles de la maladie.

### **I.2.1 Manifestations cliniques**

Si les principaux symptômes sont articulaires, d'autres sites peuvent être touchés.

#### **I.2.1.1 Manifestations articulaires**

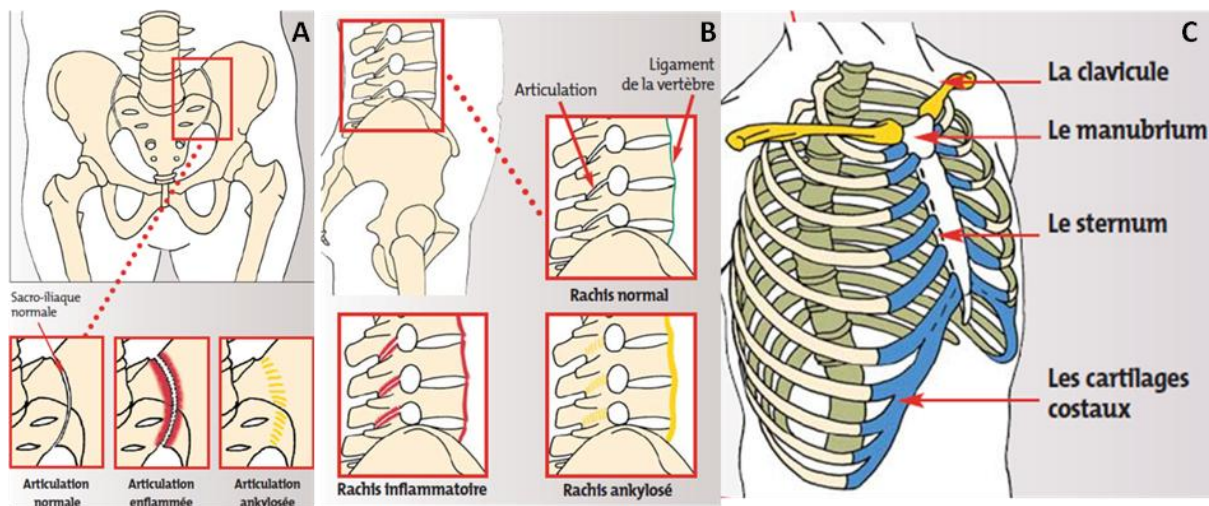
Les manifestations articulaires de la maladie se subdivisent en trois catégories : les manifestations axiales, les manifestations enthésitiques périphériques et les arthrites périphériques.

##### *Manifestations axiales*

L'atteinte axiale est caractéristique de la SpA avec trois localisations principales : les articulations sacro-iliaques, le rachis et la paroi thoracique antérieure (**Figure 1**). L'inflammation de ces sites articulaires se traduit par deux symptômes majeurs : douleur et raideur.

Les douleurs sont d'horaire inflammatoire, c'est-à-dire qu'elles réveillent le patient en deuxième partie de nuit et diminuent dans la journée avec l'activité. Les douleurs des sacro-iliaques sont à l'origine du tableau typique de lombo-fessalgie à bascule (douleur lombo-fessière alternant d'un côté à l'autre). Une irradiation à la face postérieure de la cuisse lui est fréquemment associée. Le siège de la douleur peut aussi être dorsal, cervical ou au niveau de la paroi thoracique antérieure. L'évolution se fait le plus souvent par poussées entrecoupées d'intervalles libres. Ces douleurs sont fréquemment accompagnées d'une raideur inflammatoire prédominant au réveil, de durée variable : on parle de dérouillage matinal.

La survenue d'une ankylose, conséquence d'une ossification ligamentaire progressive, constitue le risque principal de l'atteinte axiale. Au niveau sacro-iliaque, l'ankylose se traduit par la formation d'un pont osseux réunissant les deux berges de l'articulation et pouvant aboutir à sa disparition, sans entraîner toutefois d'importantes conséquences fonctionnelles. L'ankylose peut aussi survenir au niveau du rachis et de la cage thoracique et entraîner une diminution de leur souplesse.



**Figure 1 – Représentations des atteintes articulaires axiales dans la SpA ; d'après Dougados *et al.*<sup>26</sup>**  
A – articulations sacro-iliaques ; B – rachis lombaire ; C – paroi thoracique antérieure

### *Syndrome enthésitique périphérique*

L'enthèse désigne la zone fibro-cartilagineuse d'insertion des tendons, ligaments et capsules articulaires dans l'os (**Figure 2A**). Elle contribue à répartir les forces de tension sur l'ensemble de l'interface osseuse. L'enthésite, lésion inflammatoire de l'enthèse, est l'une des atteintes les plus caractéristiques de la SpA. Elle évolue en trois phases successives : une phase inflammatoire initiale, une phase érosive et enfin une phase de cicatrisation qui entraîne une fibrose puis une ossification à l'origine de la formation d'enthésophytes<sup>27</sup>.

L'atteinte la plus fréquente se situe au niveau du tendon d'Achille ou de l'aponévrose plantaire. Elle se traduit par des talalgies. Toutes les enthèses fibro-cartilagineuses peuvent être touchées, mais les enthèses périphériques les plus fréquemment atteintes sont situées au niveau de la tubérosité tibiale antérieure, du grand trochanter, de l'insertion patellaire du quadriceps et des épicondyles (**Figure 2B**)<sup>28</sup>.

### *Atteintes articulaires périphériques*

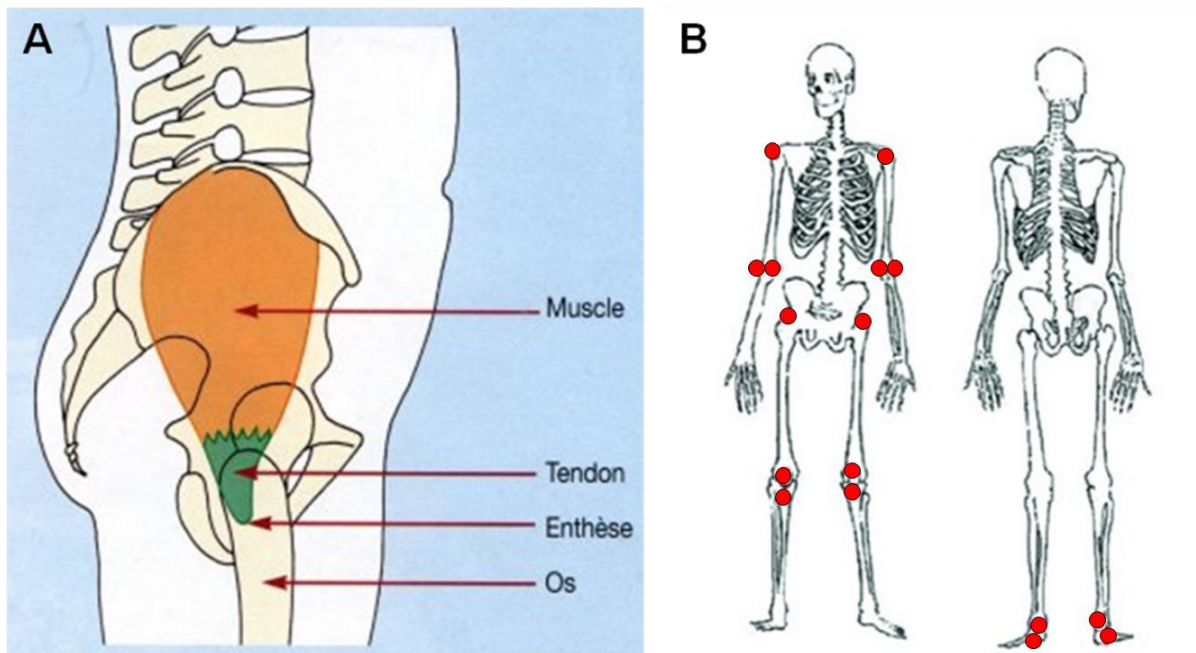
Les arthrites périphériques sont la conséquence d'une inflammation de la membrane synoviale qui enveloppe l'articulation. Cette inflammation (synovite) entraîne un épaississement de la membrane synoviale et un épanchement intra-articulaire. Parfois, la synovite peut retentir sur le cartilage et l'os sous-jacent, entraînant leur destruction : on parle alors d'arthrite érosive.

Dans la SpA, la fréquence des arthrites est évaluée entre 40 et 55%<sup>29,30</sup>. Ces arthrites se distinguent de celles observées dans la polyarthrite rhumatoïde par leur caractère fréquemment asymétrique, par l'atteinte préférentielle des grosses articulations des membres inférieurs (hanches,



genoux, chevilles) et par leur évolution le plus souvent régressive en quelques jours à quelques mois. Elles sont fréquemment associées à des atteintes enthésitiques de voisinage.

La dactylite ou « doigt en saucisse » est une manifestation clinique très évocatrice du diagnostic de SpA. Il s'agit d'une tuméfaction globale d'un ou plusieurs doigts ou orteils, douloureuse, associée à une coloration rouge foncée. Cette atteinte correspond le plus souvent à l'association de ténosynovites à des enthésites et parfois aussi à des arthrites.



**Figure 2 – Atteintes des enthèses dans la SpA ; d'après Dougados *et al.*<sup>26</sup> et Maksymowych *et al.*<sup>31</sup>**  
A – enthèse trochantérienne ; B – sites enthésitiques les plus fréquemment touchés

La SpA peut toucher d'autres sites que les articulations, témoignant du caractère systémique de cette affection.

#### 1.2.1.2 Manifestations extra-articulaires

Bien que la SpA puisse toucher la quasi-totalité des organes, les atteintes les plus fréquentes concernent l'œil, la peau et la muqueuse intestinale.

##### *Atteinte ophtalmique*

L'atteinte ophtalmique est l'atteinte extra-articulaire la plus fréquente<sup>32</sup>. Une méta-analyse publiée en 2008 estimait qu'elle concernait environ un tiers des patients souffrant de SpA<sup>33</sup>.

Il s'agit le plus souvent d'une uvéite antérieure aiguë non granulomateuse, c'est-à-dire une inflammation de la chambre antérieure de l'œil (iritis) et du corps ciliaire (cyclite), formant le segment antérieur de l'œil. Cliniquement, elle se traduit par une douleur, une rougeur oculaire avec larmoiements, photophobie et baisse d'acuité visuelle d'installation brutale.

### *Atteinte cutanée*

Le psoriasis affecte environ 1,5 à 3% de la population générale<sup>34,35</sup>, mais survient beaucoup plus souvent chez les patients atteints de SpA avec des fréquences pouvant aller de 8 à 20%<sup>29,32</sup>.

D'une manière générale, le psoriasis prédomine sur les faces d'extension des membres ainsi qu'au niveau du sacrum, du cuir chevelu et des ongles. Associé à la SpA, il n'a pas de particularité par rapport au psoriasis isolé, en dehors d'une fréquence augmentée de l'atteinte des ongles<sup>36</sup>. Il précède dans plus des deux tiers des cas les manifestations articulaires, mais apparaît parfois plusieurs années après le développement de la SpA<sup>37</sup>.

### *Atteinte digestive*

Environ 8% des patients atteints de SpA souffrent en outre d'une maladie inflammatoire chronique de l'intestin (MICI), qu'il s'agisse de la maladie de Crohn ou de la rectocolite hémorragique<sup>32</sup>.

Les MICI se traduisent par une diarrhée glairo-sanglante, chronique ou évoluant en poussées aiguës, associée à de violentes douleurs abdominales. Le diagnostic repose sur l'endoscopie digestive avec étude histologique de biopsies. En dehors de toute MICI, on observe fréquemment des lésions histologiques inflammatoires aiguës ou chroniques chez les patients atteints de SpA (>60% des SpA en endoscopie) parfois même en l'absence de toute symptomatologie clinique<sup>38</sup>.

### *Autres atteintes*

D'autres organes peuvent être touchés, mais de façon beaucoup plus rare<sup>39,40</sup>. Au niveau génital, on peut observer des uréthrites aseptiques, qui font partie du tableau de l'arthrite réactionnelle, et plus rarement des prostatites ou des orchio-épididymites également aseptiques. L'atteinte cardiaque est peu fréquente et se manifeste par des troubles de la conduction auriculo-ventriculaire et par des valvulopathies aortiques ou mitrales. L'atteinte restrictive pulmonaire est une conséquence tardive de l'ankylose de la paroi thoracique : elle peut exceptionnellement se compliquer d'un emphysème apical (« maladie bulleuse des sommets »).

La diversité des manifestations observées fait qu'il est parfois difficile d'établir un diagnostic de SpA uniquement sur des critères cliniques. Bien qu'à ce jour, il n'existe aucun examen complémentaire permettant à lui seul de faire le diagnostic de SpA, certains peuvent toutefois aider au diagnostic.

## I.2.2 Examens complémentaires

En dehors du typage HLA-B27, la place des examens biologiques reste marginale dans le diagnostic de SpA. Au contraire, l'importance accordée aux examens d'imagerie comme l'imagerie par résonance magnétique (IRM) ou l'échographie est en constante augmentation.

### I.2.2.1 Examens biologiques

#### *Typage HLA-B27*

La fréquence de l'allèle HLA-B27 est beaucoup plus élevée chez les patients atteints de SpA que dans la population générale. En effet, le risque d'être atteint d'une SpA est multiplié par 40 chez les patients porteurs de l'allèle HLA-B27<sup>4</sup>. La recherche du HLA-B27 peut ainsi constituer une aide au diagnostic en cas de contexte clinique évocateur.

Cependant, la présence du HLA-B27 n'est ni suffisante, ni nécessaire au développement de la maladie et ne doit donc pas être considérée comme un marqueur diagnostique<sup>41,42</sup>.

#### *Protéines de l'inflammation*

De façon inattendue pour une maladie inflammatoire, la vitesse de sédimentation et la protéine C-réactive (CRP) ne sont supérieures à la normale que chez la moitié des patients atteints de SpA<sup>43</sup>. Par conséquent, ces deux marqueurs sériques ne peuvent pas être considérés comme des marqueurs diagnostiques valables<sup>41</sup>. Ils peuvent néanmoins être utiles dans l'évaluation de l'activité de la maladie et ont d'ailleurs été intégrés dans un nouvel indice d'évaluation de cette activité, l'Ankylosing Spondylitis Disease Activity Score (ASDAS)<sup>44</sup>.

### I.2.2.2 Imagerie

L'atteinte structurale de la maladie peut être mise en évidence par la radiologie conventionnelle ou la tomodensitométrie (TDM). D'autres examens, comme l'IRM ou l'échographie des enthèses, permettent également une évaluation plus fonctionnelle ouvrant la voie à un diagnostic plus précoce.

#### *Radiologie conventionnelle*

Les radiographies standard peuvent être utiles au diagnostic en cas d'atteintes suffisamment évocatrices. Selon les critères de New York modifiés, la présence d'une sacro-iliite confirmée (**Tableau 2**) est même indispensable au diagnostic de SA<sup>45</sup>.

**Tableau 2 – Critères de New York modifiés de SA ; d’après van der Linden *et al.*<sup>45</sup>**

Critère radiologique (bassin de face)	Critères cliniques
<ul style="list-style-type: none"> <li>• Stade 0 : aspect normal de l’articulation sacro-iliaque</li> <li>• Stade 1 : diminution focalisée de l’interligne articulaire, aspect de pseudo-élargissement, début de sclérose</li> <li>• Stade 2 : déminéralisation sous-chondrale, début d’érosion avec aspect flou et irrégulier de l’articulation</li> <li>• Stade 3 : érosions franches, sclérose importantes, condensation des berges articulaires, diminution de l’interligne</li> <li>• Stade 4 : fusion des berges articulaires avec ankylose de l’articulation</li> </ul>	<ul style="list-style-type: none"> <li>• Lomбалgie depuis au moins trois mois, améliorée par l’activité physique et non soulagée par le repos</li> <li>• Limitation de la mobilité du rachis lombaire dans le plan sagittal et frontal</li> <li>• Ampliation thoracique diminuée par rapport aux valeurs de sujets de même âge et de même sexe</li> </ul>

➤ Le diagnostic de spondylarthrite ankylosante est retenu si au moins un des trois critères cliniques et le critère radiologique (sacroiliite  $\geq$  stade III unilatéral ou sacroiliite  $\geq$  stade 2 bilatéral) sont remplis

Les atteintes radiologiques typiques sont les conséquences d’érosions osseuses ou d’ossification, que ce soit au niveau sacro-iliaque, rachidien ou enthésitique. Néanmoins, il est à noter que la plupart de ces modifications structurales radiologiques apparaissent tardivement et que, par conséquent, la radiographie standard n’est pas un examen de choix pour le diagnostic précoce de SpA<sup>46</sup>.

### *Tomodensitométrie*

Le scanner des articulations sacro-iliaques peut se révéler utile dans la détection d’anomalies structurales sacro-iliaques trop fines pour être encore observées sur les radiographies standards<sup>47</sup>. Toutefois, certaines anomalies fines sont difficiles d’interprétation car elles peuvent se rencontrer chez des sujets sains.

### *Imagerie par résonance magnétique*

Depuis le début des années 1990, l’IRM des articulations sacro-iliaques ou du rachis est de plus en plus utilisée pour le diagnostic de la SpA. De nombreuses études ont en effet montré que, grâce à cette technique, il était possible d’observer des anomalies bien des années avant que celles-ci ne soient détectables sur des radiographies standards<sup>48</sup>. Cette particularité en fait aujourd’hui l’un des outils privilégiés dans le diagnostic précoce de la SpA, à tel point que l’IRM des sacro-iliaques est désormais intégrée dans les nouveaux critères diagnostiques de la maladie établis par l’ Assessment of SpondyloArthritis international Society (ASAS)<sup>49</sup>.

Les anomalies observées sont de deux types : inflammatoires (œdème osseux, synovite, enthésite ou capsulite) ou structurales (sclérose sous-chondrale, érosions, dégénérescence graisseuse ou syndesmophytes)<sup>50,51</sup>.

### *Échographie des enthèses*

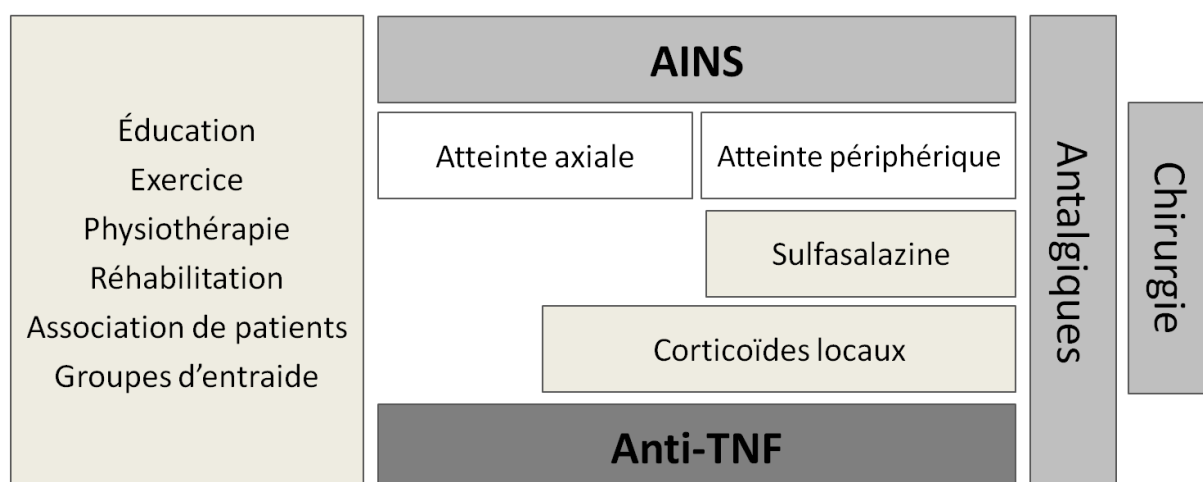
L'échographie est capable d'analyser finement la morphologie des enthèses et de visualiser des lésions de petite taille. Le couplage aux techniques de Doppler puissance permet le repérage d'une hypervascularisation locale reflétant l'inflammation<sup>52</sup>. L'échographie est plus sensible que l'examen clinique et que d'autres modalités d'imagerie dans la détection des enthésites périphériques<sup>53,54</sup>. Sa place dans le diagnostic reste toutefois à préciser.

Le diagnostic de SpA repose sur un faisceau d'arguments cliniques et paracliniques. L'objectif actuel est de détecter la maladie le plus précocement possible afin de pouvoir initier rapidement un traitement adapté.

### **I.2.3 Traitement**

En l'absence de traitement curatif, la stratégie thérapeutique actuelle est essentiellement symptomatique, avec comme principaux objectifs la lutte contre la douleur et la prévention de l'enraidissement.

Le traitement médicamenteux s'inscrit dans une prise en charge globale incluant la kinésithérapie, l'utilisation de traitements locaux et, dans de rares indications, la chirurgie. L'information et l'éducation du patient tiennent aussi une place prépondérante. L'ASAS et la European League Against Rheumatism (EULAR) ont établi en 2010 des recommandations pour la prise en charge thérapeutique des SpA (**Figure 3**)<sup>55</sup>.



**Figure 3 – Recommandations ASAS/EULAR de prise en charge de la SA ; d'après Zochling *et al.*<sup>56</sup>**

### *1.2.3.1 Traitements médicamenteux*

Les traitements médicamenteux reposent essentiellement sur les anti-inflammatoires non stéroïdiens (AINS). Les corticoïdes, tout comme les traitements de fond dits « conventionnels », n'ont pas fait la preuve d'une réelle efficacité. En cas de forme résistante aux AINS, la seule alternative actuelle est la prescription d'anti-Tumor Necrosis Factor  $\alpha$  (anti-TNF $\alpha$ ). D'autres traitements sont toutefois à l'étude.

#### *Anti-inflammatoires non stéroïdiens*

Le traitement de première intention de la SpA repose sur les AINS. Ils sont efficaces à la fois sur la douleur et sur la raideur chez une majorité des patients<sup>57</sup>. Leur efficacité s'observe le plus souvent dans les 48 à 72 heures suivant leur introduction. D'une manière générale, tous les AINS ont une efficacité similaire dans la SpA, s'ils sont prescrits à la posologie adaptée<sup>58</sup>.

Néanmoins, il existe une grande variabilité interindividuelle de l'efficacité en fonction du type d'AINS : en cas d'échec d'un AINS, il est donc recommandé d'en essayer au moins un autre avant de conclure à leur inefficacité. Plus récemment, plusieurs études ont suggéré que les AINS pourraient retarder la progression structurale de la maladie<sup>59</sup>.

#### *Corticoïdes*

La place des corticoïdes par voie générale est marginale dans la prise en charge de la SpA. Une étude récente de faible effectif a toutefois conclu à une supériorité d'une corticothérapie à 50 mg/jour par rapport à un placebo<sup>60</sup>.

#### *Traitements de fond conventionnels*

Les traitements de fond conventionnels couramment utilisés pour la PR, tels que le méthotrexate, la sulfasalazine ou le léflunomide, s'adressent parfois aux patients qui ne répondent pas suffisamment aux AINS et/ou aux gestes locaux.

Cependant, leur efficacité n'a pas été démontrée dans les formes axiales de la maladie<sup>61,62</sup>. Ils sont encore parfois administrés pour lutter contre les manifestations périphériques, mais le niveau de preuve de leur efficacité reste faible<sup>63</sup>.

#### *Anti-tumor necrosis factor $\alpha$*

La mise sur le marché des anti-TNF $\alpha$  a marqué un tournant majeur dans la prise en charge de la SpA. Leur utilisation reposait initialement sur la mise en évidence de taux élevés d'acide ribonucléique messager (ARNm) du TNF $\alpha$  dans des prélèvements de sacro-iliaques de patients

atteints de SA<sup>64</sup>, ainsi que sur l'efficacité démontrée de l'infliximab dans la maladie de Crohn<sup>65</sup>. Les premiers essais publiés en 2000 avec l'infliximab ont montré sa grande efficacité dans la SA<sup>66,67</sup>.

Actuellement, cinq anti-TNF $\alpha$  possèdent en France une autorisation de mise sur le marché pour le traitement de la SA : l'infliximab, l'adalimumab, l'éta nercept, le golimumab et le certolizumab pegol. Si ces produits diffèrent par leur structure et par leur mode d'administration, ils semblent avoir une efficacité similaire avec 40 à 50% des patients présentant une amélioration d'au moins 40%<sup>68-72</sup>. Leurs effets sur la progression structurale de la maladie restent quant à eux controversés : en contradiction avec les premières études publiées sur le sujet<sup>73-75</sup>, plusieurs études récentes suggèrent en effet un ralentissement de la progression structurale chez les patients traités par anti-TNF $\alpha$  au long cours<sup>76,77</sup>.

### *Traitements en développement*

À l'heure actuelle, seuls les anti-TNF $\alpha$  sont indiqués en traitement de deuxième ligne chez les patients ne répondant pas aux AINS. En cas d'échec ou d'échappement à ces traitements, les possibilités thérapeutiques sont donc limitées. En effet, plusieurs biothérapies autres que les anti-TNF $\alpha$ , efficaces dans la PR, se sont révélées décevantes dans la SA : blocage de l'IL-1 (anakinra), anti-CD20 (rituximab) et anti-CTLA4 (abatacept), anti-IL-6 (tocilizumab et sarilumab)<sup>78-82</sup>.

Certaines pistes thérapeutiques actuellement en développement semblent prometteuses comme le blocage de l'IL-17 par le secukinumab<sup>83</sup>, le blocage de la sous-unité p40 commune aux récepteurs de l'IL-12 et de l'IL-23 (ustekinumab) ou encore l'inhibition de certaines enzymes (apremilast, tofacitinib, fostamatinib)<sup>84-87</sup>. Les résultats des essais thérapeutiques de large envergure en cours sont néanmoins attendus avant de pouvoir conclure à leur efficacité.

#### 1.2.3.2 Traitements physiques

Les traitements physiques, tels que la kinésithérapie, l'ergothérapie, la podologie ou l'appareillage, complètent la prise en charge médicamenteuse. Ils visent à lutter contre la douleur et l'enraidissement tout au long de l'évolution de la maladie.

L'efficacité de l'exercice physique régulier associé à la physiothérapie a été démontrée à court terme<sup>88</sup>. Cependant peu de données sont disponibles sur son efficacité à long terme.

#### 1.2.3.3 Traitements chirurgicaux

La chirurgie a une place marginale dans la prise en charge de la SpA. Les arthroplasties figurent parmi les interventions les plus couramment pratiquées, notamment au niveau de la hanche. La chirurgie du rachis, mise en œuvre pour corriger les déformations, se pratique beaucoup plus rarement.

Ces dernières années, le diagnostic et la prise en charge de la SpA ont ainsi bénéficié d'importants progrès. Une partie de ces progrès s'explique par une meilleure caractérisation de la maladie facilitant son étude.

### **I.3. Concept unifié de spondyloarthrite**

La présentation clinique de la SpA est très polymorphe. En fonction de la combinaison des différents symptômes entre eux, on distingue plusieurs sous-types de la maladie : la SA, le rhumatisme psoriasique (RP), l'arthrite réactionnelle (ARé), le rhumatisme associé aux maladies inflammatoires chroniques de l'intestin (RhMICI) et la spondylarthrite indifférenciée (USpA). Cette distinction a toutefois tendance à disparaître au profit d'une approche unifiée et des critères de classification communs ont été établis.

#### **I.3.1 Les sous-types de spondyloarthrite**

##### **I.3.1.1 Spondylarthrite ankylosante**

La SA constitue la forme la plus typique de SpA. Elle affecte préférentiellement l'homme avec un sexe ratio de 2:1 à 3:1<sup>7</sup>. Son association à l'allèle HLA-B27 est très forte, puisque plus de 90% des malades atteints de SA sont porteurs de cet allèle.

Sur le plan clinique, elle est caractérisée par la prédominance de l'atteinte axiale avec une forte tendance à évoluer vers l'ankylose articulaire. Son diagnostic nécessite d'ailleurs la mise en évidence d'une atteinte structurale évoluée des sacro-iliaques sur des radiographies standard selon les critères de New York modifiés<sup>45</sup>.

##### **I.3.1.2 Rhumatisme psoriasique**

Le RP débute en général entre 30 et 50 ans sans prédominance de sexe<sup>89</sup>. L'atteinte cutanée précède l'atteinte articulaire dans 70% des cas<sup>37</sup>. Le plus souvent, il n'y a pas de corrélation entre la sévérité de l'atteinte articulaire et celle de l'atteinte cutanée.

L'atteinte articulaire permet de distinguer deux formes de RP : la forme périphérique isolée, se rapprochant de la PR, et la forme à prédominance axiale. Les arthrites périphériques psoriasiques sont le plus souvent asymétriques et touchent préférentiellement les articulations distales (notamment les articulations interphalangiennes distales). L'atteinte axiale est proche de celle observée au cours des autres SpA et peut être associée à des arthrites, des enthésites périphériques ou des dactylites<sup>90</sup>.



### I.3.1.3 Arthrites réactionnelles

Les ARé sont des arthrites stériles survenant dans le mois qui suit une infection le plus souvent digestive ou urinaire<sup>91</sup>. L'évolution immédiate de l'ARé se fait en général vers la rémission spontanée en trois à 12 mois. Dans la plupart des cas, il s'agit d'une guérison définitive. Toutefois, dans 20 à 50% des cas, on note la survenue de récurrences, un passage à la chronicité, voire une évolution vers une authentique SA<sup>92</sup>.

Le tableau clinique le plus fréquent est celui d'une oligoarthritis asymétrique prédominant aux membres inférieurs, parfois associée à des enthésites ou des dactylites. Le début est souvent aigu, parfois associé à de la fièvre. Des manifestations extra-articulaires, comme une conjonctivite, une kératodermie blennorragique ou une urétrite aseptique, s'observent parfois.

### I.3.1.4 Arthrites associées aux maladies inflammatoires chroniques de l'intestin

Les MICI s'accompagnent fréquemment de manifestations articulaires de SpA<sup>93,94</sup>. L'atteinte articulaire associée aux MICI est caractérisée par la fréquence des arthrites périphériques et la moindre association au HLA-B27.

### I.3.1.5 Spondylarthrites indifférenciées

Les USpA regroupent l'ensemble des formes de SpA ne s'intégrant pas dans l'un des sous-types précédents. Elles évoluent parfois vers une SA ou plus rarement vers un autre sous-type de SpA. Toutefois, dans un certain nombre de cas, elles restent « indifférenciées » malgré plusieurs années de suivi<sup>95</sup>. Les USpA peuvent donc correspondre soit à des formes précoces de la maladie, soit dans certains cas à un sous-type particulier de SpA.

Il apparaît ainsi que chacun des sous-types de SpA dispose de caractéristiques propres mais de nombreux arguments justifient leur regroupement en une seule entité.

## I.3.2 La spondyloarthrite, un regroupement justifié

L'individualisation du concept de spondyloarthrite est relativement récente. En effet, il a fallu attendre 1957 pour que l'International League Against Rheumatism décrive pour la première fois un groupe de « polyarthrites rhumatoïdes atypiques » incluant la SA, le RP et les ARé<sup>96</sup>. En 1974, Moll *et al.* introduisent le concept de « spondarthrite séronégative »<sup>97</sup>.

Les arguments en faveur de ce regroupement sont à la fois cliniques, épidémiologiques et génétiques. Sur le plan clinique, certaines manifestations sont fréquentes dans l'ensemble des sous-types telles que les spondylites, la sacro-iliite, les enthésites, les uvéites, le psoriasis, ou les maladies inflammatoires chroniques de l'intestin. L'unification du concept de SpA se justifie également d'un

point de vue épidémiologique du fait de l'observation fréquente de formes consécutives (par exemple, SA « secondaire » à un RP, une ARé ou un RhMICI), ainsi que de la coexistence possible de plusieurs sous-types au sein d'une même famille. Enfin, ces cas familiaux témoignent d'un terrain génétique prédisposant partagé par ces différentes formes. La forte association de l'antigène HLA-B27, initialement décrite avec la SA<sup>98,99</sup> puis confirmée pour l'ensemble des sous-types<sup>100</sup>, renforce ce concept de terrain prédisposant partagé.

L'ensemble de ces arguments soutient l'idée que les différents sous-types de SpA ne sont pas des entités individualisées, mais bien des variations phénotypiques d'une seule et même affection. Afin de prendre en compte ce nouveau concept unifié de SpA, il a fallu établir des critères de classification propres à l'ensemble des sous-types et non plus spécifiques à chacun d'entre eux.

### I.3.3 Critères de classification de la spondyloarthrite

Les jeux de critères détaillés ci-dessous ont été établis en tant que critères de classification, c'est-à-dire qu'ils ont comme principal objectif d'harmoniser des groupes de patients atteints d'une maladie donnée afin de les utiliser en recherche clinique et fondamentale. Il faut ainsi les différencier des critères diagnostiques qui permettent de déterminer à l'échelle d'un individu (et non d'un groupe) s'il est malade ou non.

Les critères d'Amor et de l'European Spondyloarthropathy Study Group (ESSG), développés quasi-simultanément, sont les premiers à avoir pris en compte la SpA dans sa globalité. Plus récemment, l'ASAS a établi de nouveaux critères prenant en compte les nouveautés diagnostiques comme l'IRM.

#### I.3.3.1 Critères d'Amor

Les critères d'Amor permettent plusieurs modes d'entrée dans la maladie. Ils comprennent 12 items affectés chacun d'une pondération et répartis en quatre catégories (symptômes cliniques, signes radiographiques, terrain génétique et réponse au traitement) (**Tableau 3**).

Les critères de classification sont remplis si la somme des points est égale ou supérieure à six<sup>101</sup>. Testée initialement dans les formes établies de SpA, la sensibilité de ces critères était de 98% et leur spécificité de 86%.

#### I.3.3.2 Critères de l'European Spondyloarthropathy Study Group

Les critères de l'ESSG ont été élaborés dans le but de prendre en compte les patients atteints de SpA indifférenciée. Deux modes d'entrée sont possibles (critères majeurs) : soit une rachialgie inflammatoire, soit une arthrite périphérique (**Tableau 4**)<sup>102</sup>.

Un critère majeur et un critère mineur sont nécessaires pour remplir les critères de classification de SpA. La sensibilité et la spécificité de ces critères ont été toutes deux évaluées à 87%. Une étude comparative transversale des critères de classifications d'Amor et de l'ESSG portant sur 2 228 patients a montré des performances globalement similaires<sup>103</sup>.

**Tableau 3 – Critères de classification de SpA d'Amor ; d'après Amor *et al.*<sup>101</sup>**

Critères	Score
<b>A – Symptômes cliniques (ou antécédents)</b>	
• Douleurs nocturnes lombaires/dorsales ou raideur matinale lombaire/dorsale	1
• Oligoarthritis asymétrique	2
• Fessalgie (caractère à bascule)	1
• Orteil ou doigt en saucisse	2
• Talalgie ou autre enthésite	2
• Uvéite antérieure	2
• Urétrite ou cervicite non gonococcique, moins d'un mois avant le début d'une arthrite	1
• Diarrhée aiguë, moins d'un mois avant une arthrite	1
• Psoriasis, balanite, entéropathie inflammatoire chronique	2
<b>B – Signes radiographiques</b>	
• Sacroiliite (grade 2 bilatérale ou grade 3 et 4)	3
<b>C – Terrain génétique</b>	
• HLA-B27 positif ou antécédent familial de SpA, MICI ou psoriasis	2
<b>D – Réponse au traitement</b>	
• Amélioration des douleurs en 48h par les AINS ou rechute en 48h après leur arrêt	2

➤ Les critères de classification sont remplis si la somme des points est égale ou supérieure à six

**Tableau 4 – Critères de classification de SpA de l'ESSG ; d'après Dougados *et al.*<sup>102</sup>**

Rachialgies inflammatoires	OU	Synovites
	• Asymétriques	
	Ou	
	• Prédominant aux membres inférieurs	
<b>ET UN OU PLUS DES ÉLÉMENTS SUIVANTS</b>		
<ul style="list-style-type: none"> <li>• Antécédent familial au 1<sup>e</sup> ou 2<sup>e</sup> degré (spondylarthrite ankylosante, psoriasis, uvéite aiguë, arthrite réactionnelle, maladie inflammatoire chronique de l'intestin)</li> <li>• Psoriasis</li> <li>• Maladie inflammatoire chronique de l'intestin</li> <li>• Urétrite ou cervicite ou diarrhée aiguë moins d'un mois avant l'arthrite</li> <li>• Fessalgies à bascule</li> <li>• Enthésopathie (calcanéenne ou plantaire)</li> <li>• Sacro-iliite radiographique</li> </ul>		

**Tableau 5 – Critères de classification de SpA de l'ASAS ; d'après Rudwaleit *et al.*<sup>49,104</sup>**

Critères de spondylarthrite axiale	
Chez un patient présentant des lombalgies depuis ≥3 mois ayant débuté avant l'âge de 45 ans	
<b>Sacro-iliite + ≥1 critère de SpA</b>	<b>OU HLA-B27 + ≥2 critères de SpA</b>
<u>Critères de SpA</u> <ul style="list-style-type: none"> <li>Lombalgies inflammatoires</li> <li>Arthrite</li> <li>Enthésite</li> <li>Uvéite</li> <li>Dactylite</li> <li>Psoriasis</li> <li>MICI</li> <li>Réponse aux AINS</li> <li>Antécédent familial de SpA</li> <li>HLA-B27</li> <li>CRP élevée</li> </ul>	
<u>Sacro-iliite</u> <ul style="list-style-type: none"> <li>Inflammation compatible avec une sacro-iliite en IRM</li> </ul> Ou <ul style="list-style-type: none"> <li>Sacro-iliite radiologique répondant aux critères de New York modifiés</li> </ul>	
Critères de spondylarthrite périphérique	
Chez un patient présentant <b>uniquement</b> des manifestations périphériques	
Arthrite OU Enthésite OU Dactylite	
+	
<u>≥1 des manifestations suivantes</u> <ul style="list-style-type: none"> <li>Uvéite</li> <li>Psoriasis</li> <li>MICI</li> <li>Infection précédant les symptômes</li> <li>HLA-B27</li> <li>Sacro-iliite à l'imagerie</li> </ul>	<b>OU</b> <u>≥2 des manifestations suivantes</u> <ul style="list-style-type: none"> <li>Arthrite</li> <li>Enthésite</li> <li>Dactylite</li> <li>Lombalgies inflammatoire</li> <li>Antécédent familial de SpA</li> </ul>

### I.3.3.3 Critères de l'Assessment of SpondyloArthritis international Society

L'une des principales limites des critères d'Amor et de l'ESSG est leur faible capacité à reconnaître des formes précoces de SpA, notamment les formes sans atteinte radiographique. En 2009, le groupe ASAS a proposé de nouveaux critères de classification de SpA (**Tableau 5**)<sup>49,104</sup>.

Ces critères se distinguent des critères pré-existants sur deux principaux points. Tout d'abord, ils séparent les formes axiales de SpA des formes périphériques avec deux jeux de critères

distincts. Ensuite, la prise en compte pour la première fois de l'IRM comme outil de détection de la sacro-iliite devrait aider à établir plus précocement le diagnostic de SpA.

La sensibilité de ces critères est de 82,9% et leur spécificité de 84,4% pour les formes axiales. Les critères de classification des SpA périphériques font également preuve de bonnes performances avec une sensibilité de 77,8% et une spécificité de 82,9%.

La diversité de la présentation clinique de la SpA rend donc difficile à la fois son diagnostic et sa prise en charge thérapeutique. Pour progresser dans ces deux domaines, il est indispensable d'appréhender au mieux les mécanismes à l'origine de la maladie. La présence de cas familiaux de SpA et l'importance du HLA-B27 dans la survenue de la maladie orientent vers son origine génétique. Toutefois, la survenue d'une SpA n'est pas uniquement déterminée par la présence du HLA-B27 et on peut considérer qu'il s'agit d'une maladie complexe impliquant plusieurs facteurs de susceptibilité à la fois génétiques et environnementaux. Avant de détailler les connaissances actuelles sur les gènes impliqués dans la SpA, il est nécessaire de revenir sur les méthodes permettant d'identifier des facteurs génétiques de susceptibilité à une maladie complexe.

## CHAPITRE 2 - MÉTHODOLOGIE DE LA GÉNÉTIQUE DES MALADIES COMPLEXES

---

Les facteurs de susceptibilité à l'origine des maladies sont génétiques ou environnementaux. Le poids respectif de ces deux types de facteur varie cependant en fonction de la pathologie étudiée. Schématiquement, on observe d'un côté les maladies dites héréditaires, ou mendéliennes, et de l'autre les maladies dites environnementales. Entre ces deux extrêmes, il existe un continuum de pathologies, dites complexes ou multifactorielles, impliquant à la fois des facteurs génétiques et environnementaux.

Les études d'épidémiologie génétique se sont longtemps restreintes aux maladies mendéliennes. Il s'agit de maladies rares, avec une composante héréditaire marquée. Elles sont le plus souvent liées à la mutation d'un seul gène. À l'opposé, la survenue de maladie complexe est favorisée par des facteurs génétiques pouvant interagir entre eux ou avec des facteurs environnementaux.

Par conséquent, l'identification des gènes associés aux maladies complexes a nécessité le développement d'une méthodologie spécifique. Celle-ci s'appuie sur l'utilisation de marqueurs génétiques dans deux grands types d'analyse : les analyses de liaison et les analyses d'association. Au-delà des variations de la séquence génétique, l'identification de gènes de susceptibilité dans les maladies complexes peut également nécessiter l'étude de leur expression.

### **II.1. Caractérisation de la variabilité génétique humaine**

La séquence du génome humain ne diffère que de 0,1% en moyenne entre deux individus non apparentés<sup>105</sup>. Les variations interindividuelles de la séquence d'acide désoxyribo-nucléique (ADN) sont utilisées comme marqueurs génétiques. En effet, ces fragments d'ADN polymorphes définissent chacun une position précise et unique au niveau du génome.

Il existe plusieurs catégories de marqueurs génétiques. Les plus couramment utilisés sont les marqueurs microsatellites et les polymorphismes mono-nucléotidiques (Single Nucleotide Polymorphism ou SNP). Un intérêt croissant est aussi porté aux variations du nombre de copies (Copy Number Variation ou CNV).

#### **II.1.1 Microsatellites**

Les microsatellites sont de courtes séquences d'ADN formées par la répétition en tandem et en nombre variable  $n$  d'un même motif composé de deux à dix nucléotides. Pour un microsatellite donné, chaque valeur de  $n$  définit un allèle (**Figure 4A**).

Les microsatellites présentent plusieurs caractéristiques qui en font de bons marqueurs génétiques : ils sont très polymorphes (c'est-à-dire qu'ils possèdent un grand nombre d'allèles), leur distribution est relativement homogène sur l'ensemble du génome et ils sont facilement détectables par la technique de réaction de polymérisation en chaîne (Polymerase Chain Reaction ou PCR). Pour ces raisons, ils ont été largement utilisés depuis leur identification en 1989<sup>106,107</sup>, notamment pour la cartographie du génome<sup>108</sup> ou pour l'identification des gènes impliqués dans les maladies mendéliennes<sup>109</sup>.

Leur utilisation s'est néanmoins restreinte ces dernières années au profit des SNPs, plus denses et plus faciles à génotyper par des techniques à haut-débit.

### II.1.2 Polymorphismes mono-nucléotidiques

Les SNPs correspondent à des variations généralement bi-alléliques d'un seul nucléotide dans la séquence d'ADN (**Figure 4B**). Près de 15 millions d'entre eux ont été identifiés à ce jour, ce qui représente plus de 90% des variations observées chez l'être humain<sup>110</sup>.

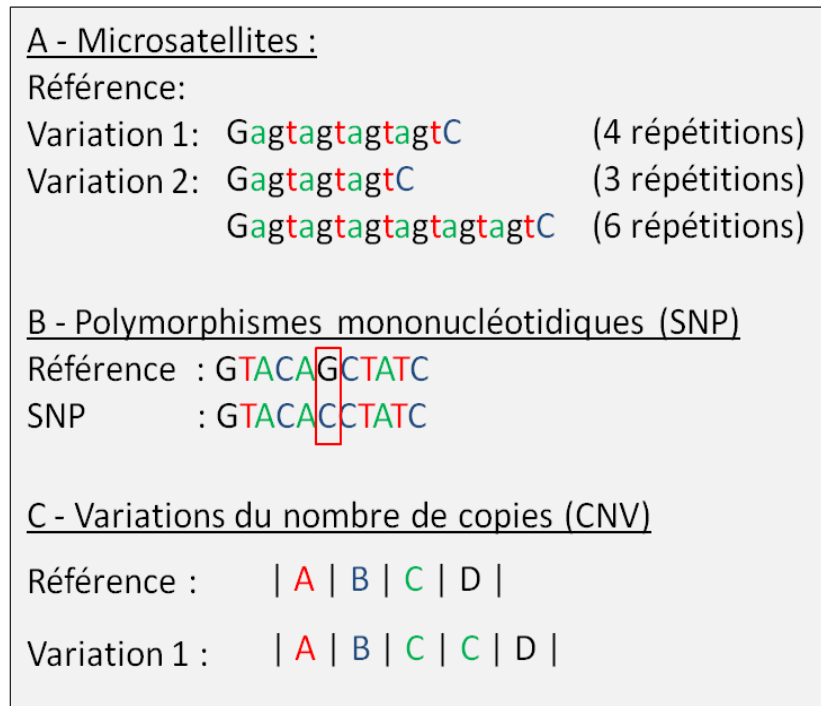
Bien qu'ils soient moins polymorphes que les marqueurs microsatellites, les SNPs présentent plusieurs avantages par rapport à ces derniers, ce qui explique leur large utilisation actuelle pour l'étude génétique des maladies complexes. Leur premier avantage est leur très grande fréquence avec une répartition relativement homogène sur tout le génome. En outre, ils sont situés aussi bien dans des régions régulatrices ou codantes que non codantes et peuvent donc avoir des conséquences fonctionnelles directes (en modifiant par exemple la structure d'une protéine pour les SNPs situés dans une région codante ou son niveau d'expression pour ceux situés dans les régions régulatrices)<sup>111</sup>. Enfin, les progrès des techniques de biologie moléculaire permettent le génotypage simultané à faible coût d'un très grand nombre de SNPs, à l'aide de puces à ADN.

### II.1.3 Variations du nombre de copies

Un CNV correspond à tout segment d'ADN d'une kilobase (kb) ou plus présentant un nombre de copies variable entre les individus d'une même espèce<sup>112</sup>. Ils peuvent être la conséquence d'événements de duplication, de délétion, d'insertion ou de translocation et présenter différents degrés de complexité (**Figure 4C**)<sup>113</sup>. La première carte des CNVs du génome humain a été publiée en 2006 : elle a permis d'établir qu'au moins 360 mégabases (Mb) d'ADN étaient concernées par ce type de polymorphisme, soit plus de 12% du génome humain<sup>113</sup>.

Les CNVs ont été peu étudiés jusqu'à il y a quelques années, car la résolution obtenue par les techniques traditionnellement employées était trop faible pour détecter la plupart d'entre eux. Le développement du génotypage haut-débit et plus récemment du séquençage dit de « nouvelle génération » ont cependant facilité leur identification, y compris des petits CNVs<sup>114</sup>. Ceux-ci ont un

impact fonctionnel potentiel et sont impliqués dans de nombreuses pathologies, qu'elles soient mendéliennes ou complexes<sup>115</sup>.



**Figure 4 – Différents types de marqueur génétique**

Les marqueurs génétiques, et particulièrement les SNPs, sont utilisés dans deux grands types d'étude génétique : les analyses de liaison et les analyses d'association.

## **II.2. Analyses de liaison génétique**

On considère que deux locus sont liés s'ils sont transmis ensemble d'un parent à son enfant plus souvent qu'on ne l'attendrait en cas d'indépendance. Autrement dit, si, durant la méiose, une recombinaison entre deux locus survient à une fréquence inférieure à 50%, alors ces deux locus sont dits liés<sup>116</sup>. Les analyses de liaison génétique utilisent ce concept pour tester l'indépendance de transmission familiale entre des marqueurs génétiques et le locus responsable de la maladie (locus-maladie) (**Figure 5**).

On distingue deux principaux types d'analyse de liaison selon qu'elles reposent ou non sur la possibilité de définir au préalable un modèle de transmission de la maladie : les analyses paramétriques et les analyses non paramétriques.

### **II.2.1 Analyses de liaison paramétriques**

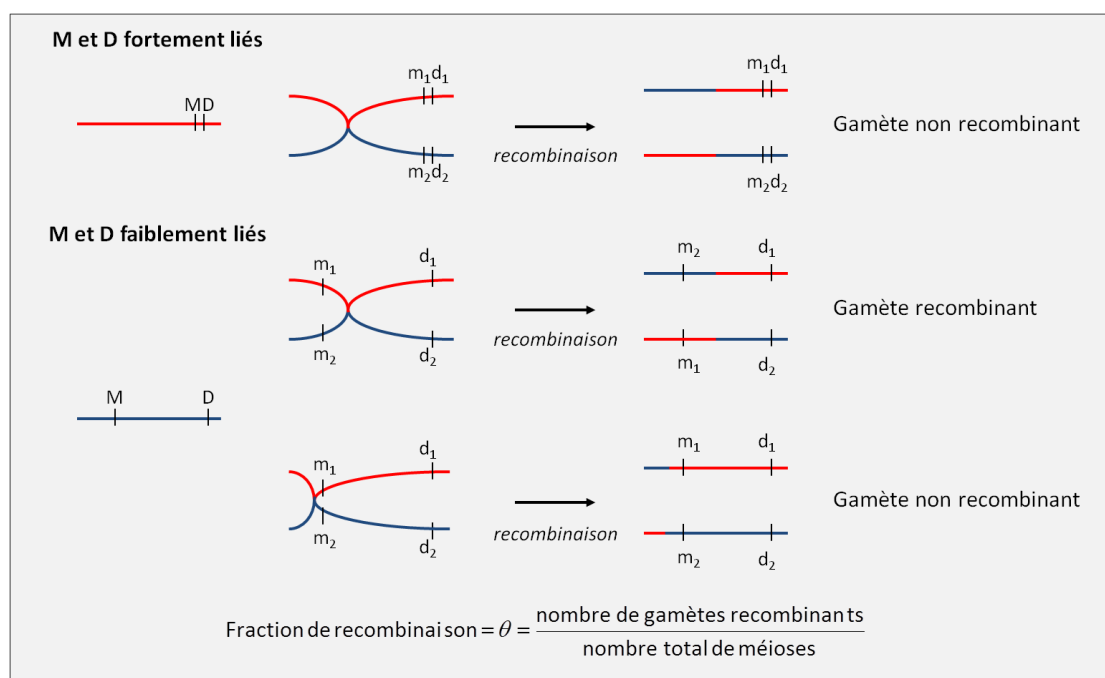
Les analyses de liaison paramétriques sont basées sur le calcul d'un score de liaison, le logarithm of odds score (LOD score). Initialement destinées à l'identification de gènes de



susceptibilité impliqués dans les maladies mendéliennes, elles ont été secondairement adaptées pour une utilisation dans les maladies complexes.

### II.2.1.1 Calcul du LOD score

En génotypant des marqueurs répartis sur l'ensemble du génome et en étudiant la co-transmission de ces marqueurs avec un trait phénotypique, il est possible de localiser une région génétique, ou locus, spécifique à la maladie. Les analyses de liaison paramétriques se basent sur l'estimation de la fraction de recombinaison  $\theta$ , c'est-à-dire la proportion de recombinaison entre deux locus lors de la méiose (**Figure 5**).



**Figure 5 – Liaison génétique et fraction de recombinaison**

M : marqueur ; D : Disease = locus-maladie

En 1947, Haldane et Smith ont proposé de calculer un rapport de vraisemblance (likelihood ratio) pour estimer la liaison entre deux locus<sup>117</sup>. Ce rapport de vraisemblance est calculé en prenant en compte la fraction  $\theta$ , la fréquence de l'allèle morbide et les taux de pénétrance de la maladie en fonction du génotype testé (**Figure 6A**). Il est ensuite maximisé (maximum likelihood ratio ou MLR) : on conserve la valeur la plus élevée obtenue en faisant varier  $\theta$  de 0 à  $\frac{1}{2}$  (**Figure 6B**)<sup>117,118</sup>. Pour simplifier son utilisation, Morton a proposé en 1955 d'utiliser le logarithme décimal du MLR<sup>119</sup> (**Figure 6C**). Le LOD score ainsi obtenu facilite la prise en compte de l'information de plusieurs familles, puisqu'il suffit de faire la somme de leurs LOD scores respectifs pour obtenir un LOD score global.

Si au départ, le calcul du LOD score n'était possible que pour des structures familiales simples, des algorithmes ont secondairement été développés afin de permettre sa généralisation à tous types de structure<sup>120-123</sup>. Ces différents algorithmes sont à la base de plusieurs programmes informatiques comme Linkage<sup>124</sup>, GeneHunter<sup>125</sup> ou Merlin<sup>126</sup>.

<b>A</b>	<b>Likelihood ratio</b> = $\frac{\text{Vraisemblance des résultats sous l'hypothèse de liaison } (\theta < 1/2)}{\text{Vraisemblance des résultats sous l'hypothèse d'indépendance } (\theta = 1/2)}$
<b>B</b>	<b>Maximum Likelihood Ratio</b> = MLR = $\max_{0 < \theta < 1/2} (\text{Rapport de vraisemblance})$
<b>C</b>	<b>Logarithm of odds</b> = LOD = $\log(\text{MLR})$
<b>D</b>	<b>HLOD</b> ( $\alpha, \theta$ ) = $\log_{10} \left( \frac{\text{Vraisemblance des résultats sous l'hypothèse } \alpha < 1 \text{ et } \theta < 1/2}{\text{Vraisemblance des résultats sous l'hypothèse } \alpha = 1 \text{ et } \theta = 1/2} \right)$

**Figure 6 – Définition des principaux scores de liaison**

La méthode décrite précédemment est dite bi-point, c'est-à-dire que la fraction de recombinaison est calculée entre un seul marqueur génétique et le locus morbide. Pour une plus grande précision, des méthodes multi-points ont été développées permettant de localiser le locus morbide par rapport à une carte fixe de marqueurs génétiques dont la position est connue<sup>127</sup>. L'analyse multi-point est donc plus puissante, car elle prend en compte simultanément les informations données par un grand nombre de marqueurs. Elle nécessite toutefois de connaître la distance entre chacun des marqueurs utilisés et requiert des capacités informatiques beaucoup plus importantes que l'analyse bi-point.

### II.2.1.2 Prise en compte de l'hétérogénéité de locus : LOD score d'hétérogénéité

Un des principaux facteurs pouvant affecter la puissance des analyses de liaison est la présence d'une hétérogénéité de locus. Cette hétérogénéité existe quand des allèles morbides situés sur plusieurs locus indépendants sont à l'origine du même phénotype. Dans ces cas, si une analyse de liaison est réalisée avec un marqueur situé à proximité d'un de ces locus, la fraction de recombinaison varie selon les familles. L'analyse combinée de ces deux types de familles risque alors d'empêcher la détection du signal de liaison alors que celui-ci existe bien<sup>128</sup>.

Un des moyens de détecter la liaison en cas d'hétérogénéité de locus consiste en l'introduction d'un paramètre supplémentaire pris en compte lors de l'analyse : le coefficient  $\alpha$ ,

correspondant à la proportion de familles qui possède la forme génétique de la maladie liée au marqueur d'intérêt<sup>129</sup>. Ce test, appelé « admixture test », permet de calculer le LOD score d'hétérogénéité ou HLOD<sup>130,131</sup> (Figure 6D).

#### II.2.1.3 Interprétation des résultats

Classiquement, l'interprétation du LOD score se fait selon les règles suivantes<sup>116</sup> :

- LOD score  $\geq 3$  : liaison entre le marqueur testé et le locus maladie,
- LOD score  $< -2$  : absence de liaison,
- LOD entre  $-2$  et  $3$  : impossible de conclure à la présence ou non d'une liaison.

Afin de prendre en compte le caractère imprécis de l'estimation de la fraction de recombinaison, il est possible de déterminer, en cas de liaison significative, l'intervalle de support. Cet intervalle correspond à la région située de part et d'autre du pic de liaison pour laquelle le LOD score dépasse un seuil correspondant au LOD score maximum moins une unité<sup>132</sup>.

Les seuils de significativité ci-dessus ont été déterminés pour des analyses n'utilisant qu'un faible nombre de marqueurs. Avec l'avènement des études génome-entier et l'augmentation du nombre de marqueurs, il est apparu indispensable de prendre en compte la multiplicité des tests. En effet, une même hypothèse nulle d'absence de liaison avec la maladie est testée pour chaque marqueur. Le risque de faux-positif augmente donc, ce qui doit être pris en compte lors de la détermination du seuil de significativité (seuil  $\alpha$  de l'erreur de type I). Lander & Kruglyak ont ainsi montré qu'en cas d'analyse de liaison génome-entier, un LOD score de 3 correspond à un risque alpha de seulement 0,09. Pour ramener ce risque au seuil couramment utilisé de 0,05, le seuil de LOD score doit être fixé à 3,3<sup>133</sup>.

#### II.2.1.4 Analyses paramétriques et maladies complexes

Bien que la nécessité de spécifier un modèle de transmission limite a priori l'utilisation des analyses paramétriques aux maladies mendéliennes, certains auteurs ont suggéré qu'elles pouvaient également être utilisées pour l'étude des maladies complexes, sous réserve de quelques précautions. Effectivement, même si une erreur de modèle risque d'entraîner une perte de puissance pour détecter une liaison existante<sup>134</sup>, plusieurs méthodes basées sur la maximisation du LOD score en fonction des paramètres du modèle ont été développées<sup>134,135</sup>. La puissance de ces méthodes est similaire à celle obtenue par une spécification correcte du mode de transmission<sup>136</sup>.

Toutefois, le fait de considérer plusieurs modèles génétiques différents multiplie le risque de fausse positivité et doit donc conduire à corriger les seuils de significativité<sup>137</sup>. Afin de limiter l'augmentation de ce seuil, Greenberg *et al.* ont proposé l'approche MMLS-C (corrected maximized maximum LOD score)<sup>138</sup>. Cette approche consiste à calculer les LOD scores sous deux modèles

simples, dominant et récessif, chacun avec une pénétrance arbitraire de 50%, puis de sélectionner le modèle avec lequel on obtient le LOD score le plus élevé et d'utiliser un seuil de significativité corrigé de 3,3.

Si l'analyse de liaison paramétrique est donc une option possible pour les maladies complexes, la méthode de référence dans ce type de maladie reste tout de même l'analyse de liaison non paramétrique.

## II.2.2 Analyses de liaison non paramétriques

Face à la difficulté de déterminer le mode de transmission dans les maladies complexes, des méthodes non paramétriques ont été développées. Ces méthodes se basent également sur l'étude de la co-transmission de marqueurs génétiques avec la maladie, mais ne nécessitent pas de spécifier au préalable le mode de transmission. Pour ce faire, elles utilisent la notion de partage d'allèles identiques par descendance (Identical By Descent ou IBD) chez les individus atteints (**Figure 7**). En effet, chez des apparentés partageant un trait phénotypique (par exemple, atteints de SpA), on s'attend, quel que soit le mode de transmission, à un excès d'allèles identiques par descendance dans la région du gène responsable de la maladie.

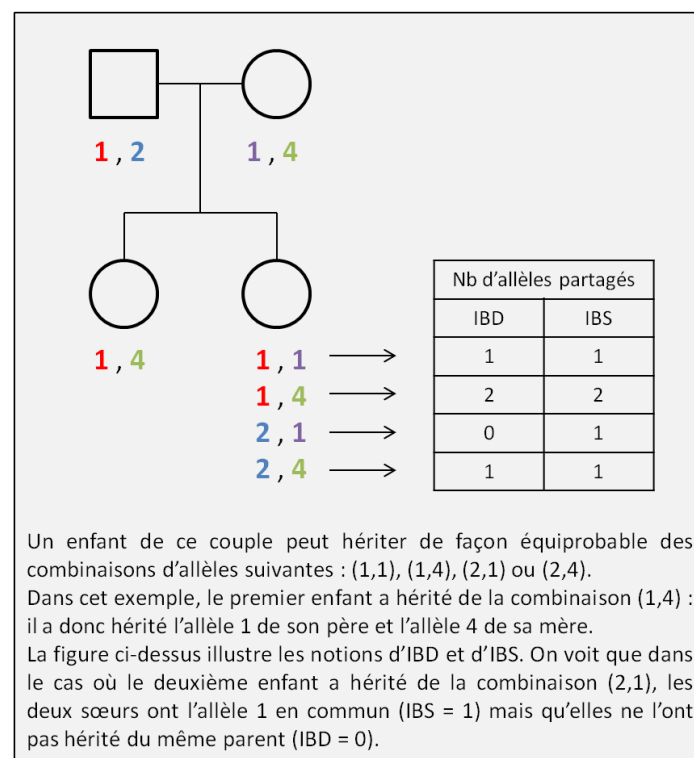
Ces méthodes ont été appliquées initialement aux paires de germains atteints (c'est-à-dire deux individus malades appartenant à la même fratrie), puis elles ont été généralisées à des familles plus complexes.

### II.2.2.1 Méthode des paires de germains atteints

L'approche la plus simple, proposée initialement par Penrose en 1935, repose sur l'étude de paires de germains atteints<sup>139</sup>. Cette méthode consiste, pour un marqueur donné, à comparer chez deux germains malades la proportion du nombre d'allèles identiques par descendance (allèles IBD) par rapport à la proportion attendue en l'absence de liaison. Plusieurs tests statistiques peuvent être appliqués pour vérifier si la proportion observée s'écarte significativement de la proportion attendue. Dans la plupart des situations, le test le plus puissant est le test dit « moyen » dans lequel la moyenne du nombre d'allèles IBD est comparée à la moyenne attendue de 1<sup>140</sup>.

En pratique, il est souvent difficile de déterminer de façon certaine le partage IBD entre paires de germains. En effet, l'absence des génotypes parentaux ou le faible polymorphisme des marqueurs peuvent rendre impossible la distinction entre IBD et identité par état (identical by state ou IBS) (**Figure 7**). Plusieurs méthodes ont été proposées pour remédier à ce problème. La plus simple est de ne conserver que les paires de germains pour lesquelles l'IBD peut être déterminée

sans ambiguïté<sup>141</sup> : il en résulte toutefois une importante perte de puissance. Une autre solution consiste à utiliser l'IBS au lieu de l'IBD<sup>142,143</sup>. L'inconvénient de cette méthode est qu'elle nécessite d'estimer les fréquences alléliques dans la population. Il est souvent difficile d'établir ces fréquences avec précision et une erreur sur leur estimation peut conduire à la détection d'une fausse liaison<sup>144</sup>. Plus récemment, Risch a proposé une méthode basée sur le calcul du maximum de vraisemblance (Maximum Likelihood Score ou MLS)<sup>145,146</sup>. Ainsi, en l'absence des génotypes parentaux, les proportions du partage IBD sont calculées en considérant tous les génotypes parentaux possibles qui seraient en accord avec les données.



**Figure 7 – Partage d'allèles entre germains - notion d'IBD et d'IBS ; d'après Dawn Teare *et al.*<sup>116</sup>**

### II.2.2.2 Extension à d'autres types de structures familiales

La méthode des paires de germains atteints a secondairement été étendue à des structures familiales plus complexes. À partir des travaux initiaux de Whittemore et Halpern<sup>147,148</sup>, Kruglyak *et al.* ont proposé en 1996 la statistique NPL (non parametric linkage) et l'ont incorporée dans le logiciel Genhunter<sup>125</sup> (**Figure 8**).

Le NPL score, qui est une méthode dite exacte, se heurte à la difficulté de détermination de l'IBD dans certaines conditions (notamment dans les cas où il y a des génotypes manquants dans la famille). Pour remédier à ce problème, Kong et Cox ont élaboré une statistique basée sur le maximum de vraisemblance. Ils ont également transformé leur statistique Z en LOD score pour

permettre entre autres de déterminer les intervalles de support<sup>149</sup>. Ces méthodes sont utilisées par plusieurs logiciels comme Allegro ou Merlin<sup>126,150</sup>.

Pour une famille  $i$  en un point  $x$  du génome, la statistique NPL est définie comme suit :

$$z_i(x) = \frac{S_i(x) - \mu_{i0}}{\sigma_{i0}}$$

où  $S_i(x)$  = IBD entre atteints dans la famille  $i$  au point  $x$   
 $\mu_{i0}$  = espérance de  $S_i(x)$  en l'absence de liaison entre le point  $x$  et la maladie  
 $\sigma_{i0}$  = variance de  $S_i(x)$  en l'absence de liaison entre le point  $x$  et la maladie

Sur un échantillon de  $N$  familles, la statistique du test global est

$$z(x) = \sum \gamma_i Z_i(x)$$

où  $\gamma_i(x)$  = coefficient de pondération modulant l'importance d'une famille dans un échantillon

**Figure 8 – Calcul de la statistique NPL (Z score)**

### II.2.2.3 Interprétation des résultats

La détermination du seuil de significativité dans les analyses de liaison non paramétriques est complexe car ce seuil varie selon le nombre de marqueurs testés, leur étendue et le type de famille étudiée. Afin d'homogénéiser l'interprétation des études de liaison, Lander et Kruglyak ont publié en 1995 une série de recommandations qui font aujourd'hui encore référence<sup>133</sup>. Ils proposent différents seuils de LOD score correspondant à une valeur  $p$  de 5% en fonction du type de paires étudiées pour des études pan-génomiques avec une forte densité de marqueurs. Ainsi le seuil de significativité du LOD score pour les études de paires de germains est fixé à 3,6. Afin d'éviter un trop grand nombre de faux négatifs, un seuil de suggestivité est proposé (correspondant à un faux positif attendu par hasard sur l'ensemble du génome) : ce seuil est de 2,2 pour les études de paires de germains. Enfin, le seuil de LOD score pour les études de réplication a été fixé à 1,18.

Cependant ces recommandations ont été critiquées pour leur caractère arbitraire<sup>151</sup>, mais aussi parce qu'elles ne prennent pas en compte la puissance intrinsèque de l'étude : il est moins probable d'obtenir un LOD score significatif si l'effectif est petit<sup>152</sup>. La simulation de données permet également d'interpréter les résultats d'une analyse non paramétrique<sup>153,154</sup>. Dans un premier temps, des données de génotypes sont simulées sous l'hypothèse d'absence de liaison, avec la même structure de familles, la même carte de marqueurs, les mêmes fréquences alléliques et les mêmes patrons de données manquantes. Ces données sont ensuite analysées afin d'obtenir la distribution des LOD scores maximaux obtenus qui pourra être comparée aux résultats obtenus.

Malheureusement, cette approche est exigeante en ressources informatiques et peut s'avérer très longue notamment si la carte de marqueurs est très dense et si les structures familiales utilisées sont complexes.

Devant la faible agrégation familiale qui caractérise la plupart des maladies complexes, de nouvelles méthodes d'association génétique ne nécessitant pas l'utilisation de familles étendues ont été développées.

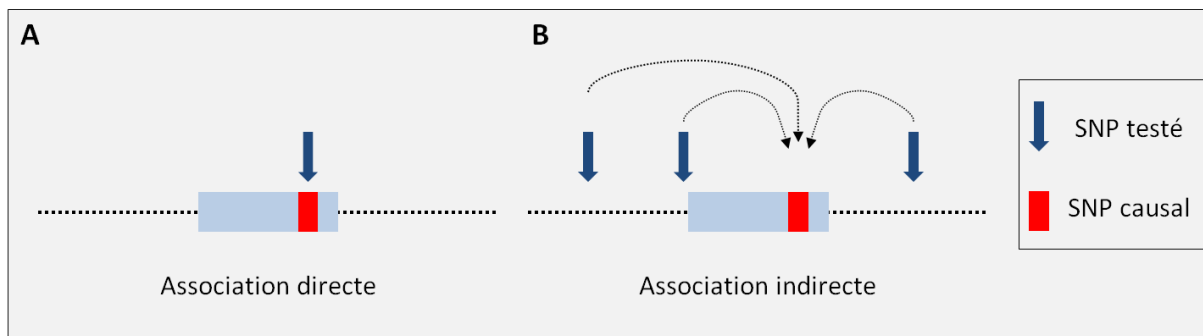
### **II.3. Analyses d'association génétique**

Les études d'association constituent une alternative ou un complément aux analyses de liaison pour l'identification de gènes de prédisposition aux maladies complexes. Elles se basent sur le principe de déséquilibre de liaison. On distingue deux grands types d'étude : les études cas-témoins et les études intra-familiales. À ces deux types d'étude, il est possible d'appliquer deux approches : gène-candidat ou génome entier. Compte-tenu de leur large utilisation durant ces dernières décennies, des recommandations ont été établies pour aider à leur conception, réalisation et interprétation.

#### **II.3.1 Déséquilibre de liaison**

Le déséquilibre de liaison (DL) se définit comme l'association non aléatoire d'allèles à des locus différents. Il correspond à la situation où la probabilité d'observer un couple d'allèles sur un chromosome n'est pas égale au produit des probabilités d'observer ces allèles individuellement<sup>155</sup>. Le plus souvent, il survient du fait de l'apparition d'un nouvel allèle par mutation. Le degré de DL entre deux locus dépend à la fois de la fraction de recombinaison  $\theta$  et du délai écoulé depuis la survenue de la mutation à l'origine d'un des deux allèles. Ainsi, le DL diminue-t-il avec l'éloignement des locus et au cours des générations successives en raison de la survenue de recombinaisons.

Ce concept a une importance fondamentale dans les analyses d'association. En effet, la mise en évidence d'une association entre un polymorphisme et une maladie ne signifie pas nécessairement que ce polymorphisme a un rôle causal dans la maladie, mais peut aussi indiquer que le polymorphisme est en DL avec le variant causal (**Figure 9**). De ce fait, il est primordial de pouvoir quantifier l'importance du DL.



**Figure 9 – Associations directe et indirecte ; d'après Hirschhorn *et al.*<sup>156</sup>**

A – le SNP testé est le SNP causal, l'association mise en évidence est dite directe ; B – les SNPs testés sont en DL avec le SNP causal. Les associations mises en évidence sont dites indirectes.

De nombreuses méthodes ont été proposées pour évaluer la force du DL. La plupart d'entre elles mesurent le DL entre des paires de marqueurs bi-alléliques. Les plus couramment utilisées sont  $|D'|$  et  $r^2$ <sup>157,158</sup> (**Figure 10**). Les deux coefficients  $|D'|$  et  $r^2$  varient de 0, correspondant à l'équilibre de liaison, à 1, correspondant à un DL complet.

Soit une paire de locus A et B présents sur le même chromosome présentant chacun 2 allèles 1 et 2. La mesure  $|D'|$  se définit comme suit :

$$|D'| = \frac{f(A_1B_1) - f(A_1) * f(B_1)}{D_{\max}}$$

où  $f$  = fréquence  
 $D_{\max}$  = valeur maximale possible du DL

Le coefficient de corrélation  $r^2$  est quant à lui calculé de la manière suivante :

$$r^2 = \frac{D^2}{f(A_1) * f(A_2) * f(B_1) * f(B_2)}$$

**Figure 10 – Calcul des coefficients de DL  $|D'|$  et  $r^2$**

La répartition du DL sur l'ensemble du génome n'est pas homogène et ne dépend pas uniquement de la distance entre les marqueurs<sup>155</sup>. Les études du DL à l'échelle du génome ont permis de montrer que le génome était divisé en régions de fort déséquilibre de liaison (blocs de DL) séparées entre elles par des points chauds de recombinaison<sup>159–161</sup>. Cette répartition en blocs de DL est particulièrement intéressante pour les études d'association génétique car elle permet de restreindre le nombre de SNPs à génotyper dans les études d'association à large échelle. Les haplotypes les plus fréquents peuvent en effet être reconstruits à l'intérieur de chaque bloc de DL. Le nombre minimal de marqueurs nécessaire pour les distinguer les uns des autres (tag-SNP ou SNP-étiquette) est ensuite déterminé<sup>162</sup>. En utilisant cette méthode, le nombre de SNPs nécessaires pour



couvrir le génome a été estimé entre 300 000 et 500 000 pour les non Africains et à 1 000 000 pour les Africains<sup>163,164</sup>.

### II.3.2 Tests d'association

Il existe deux grands types d'études d'association génétique : les études cas-témoins et les études intra-familiales.

#### II.3.2.1 Analyses cas-témoins

L'approche la plus courante consiste à comparer la fréquence de polymorphismes entre un échantillon de cas et un échantillon de témoins sains non apparentés appartenant à la même population générale. De nombreuses méthodes statistiques peuvent être utilisées, cherchant le plus souvent à mettre en évidence des différences de fréquence allélique ou génotypique entre le groupe des cas et celui des témoins. Une différence statistiquement significative correspond à deux possibilités : soit le marqueur associé est directement impliqué dans la maladie (association directe ou causale), soit il est en DL avec le variant causal (association indirecte)<sup>156</sup> (**Figure 9**).

L'analyse la plus simple consiste à tester l'association entre la maladie et chaque marqueur pris individuellement. Il existe de nombreux tests d'association allélique ou génotypique parmi lesquels les plus simples sont le test du  $\chi^2$  et le test exact de Fisher. Ces tests nécessitent que les marqueurs utilisés soient en équilibre de Hardy-Weinberg (HW) dans la population étudiée. Pour s'affranchir de ce pré-requis, le test de Cochran-Armitage (ou test de tendance), basé sur un modèle de régression linéaire, peut être utilisé<sup>165</sup>.

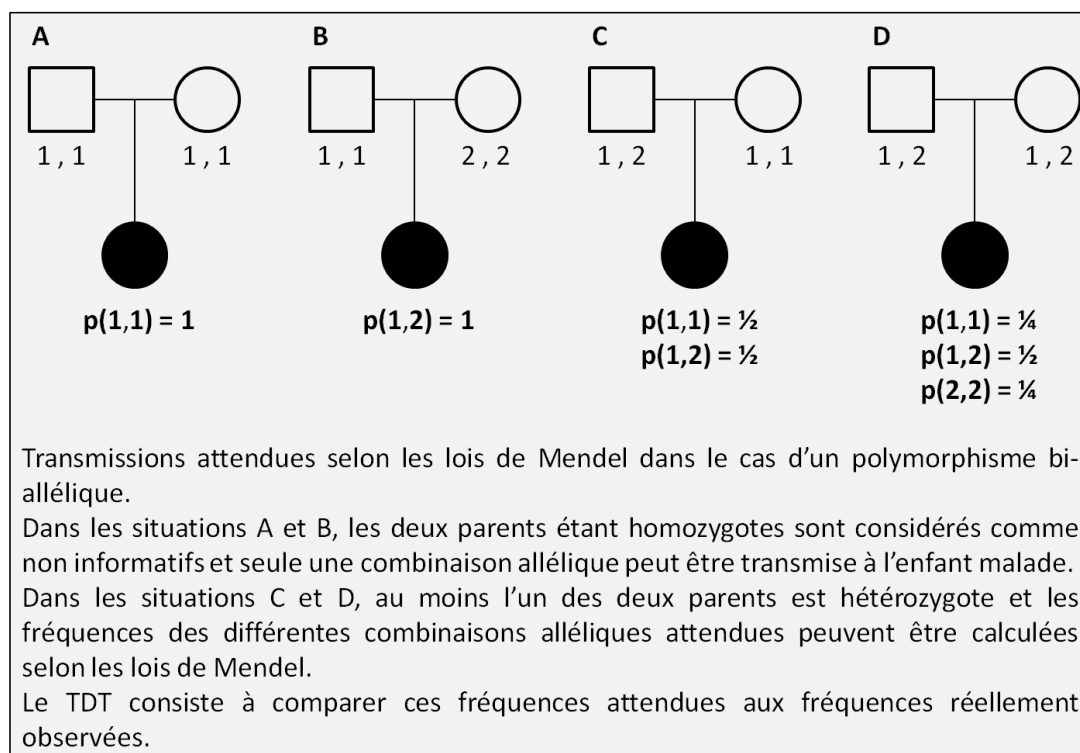
L'analyse séparée de chaque marqueur entraîne une perte d'information liée à l'absence de prise en compte du DL dans la région. Avec le nombre croissant de marqueurs génotypés dans les études d'association, des méthodes d'analyse multi-marqueurs ont été développées<sup>166</sup>. Ces analyses souffrent toutefois souvent d'un défaut de puissance en raison du nombre important d'haplotypes différents à tester. Pour remédier à ce problème, différentes approches ont été développées : certains auteurs préconisent le retrait des haplotypes rares de l'analyse, tandis que d'autres proposent des méthodes de clustering basées soit sur les similarités entre haplotypes, soit sur l'histoire évolutive des populations<sup>167</sup>.

Afin de chercher à prendre en compte simultanément les informations fournies par les analyses de liaison et d'association, des tests d'association intra-familiale ont été mis au point.

### II.3.2.2 Études d'association intra-familiale

Le trio constitue la structure familiale la plus simple dans les études d'association intra-familiale. Il est formé d'un individu malade et de ses deux parents. Le test de déséquilibre de transmission (TDT) a été conçu en 1993 par Spielman<sup>168</sup>. Il s'agit de comparer la fréquence de transmission allélique entre un parent hétérozygote et son enfant atteint à la fréquence attendue selon les lois de transmission mendéliennes (**Figure 11**). La significativité de l'association s'évalue grâce à un test de  $\chi^2$ . L'hypothèse nulle du TDT postule soit l'absence de liaison en présence d'association, soit l'absence d'association en présence de liaison, soit l'absence de liaison et d'association entre le marqueur testé et la maladie. Le fait de rejeter cette hypothèse permet donc de conclure que le marqueur testé est à la fois lié et associé à la maladie<sup>169</sup>.

De nombreuses extensions du TDT ont été élaborées pour permettre notamment l'utilisation de marqueurs multi-alléliques, de structures familiales plus complexes que des trios ou de familles dont certains parents n'ont pas été génotypés<sup>170</sup>. Des tests pouvant détecter une association en présence de liaison ont également été proposés<sup>169</sup>.



**Figure 11 – Principes du TDT**

### II.3.3 Stratégie générale

Quel que soit le type d'étude menée, deux grandes stratégies peuvent être appliquées : une approche gène-candidat ou une approche génome entier.

### II.3.3.1 Approche gène-candidat

L'approche gène-candidat est la plus intuitive pour identifier des facteurs génétiques de susceptibilité dans une maladie complexe. Elle consiste à sélectionner les polymorphismes à tester pour leur appartenance à un gène choisi en fonction de son rôle présumé dans la maladie étudiée. Le choix du gène s'effectue sur la base de connaissances physiopathologiques préalables, de l'étude de modèles animaux ou encore du rapprochement fait avec d'autres pathologies proches (comme par exemple, la possibilité de tester dans la SpA un gène de susceptibilité au psoriasis). Il peut également être guidé par des études de liaison préalables suggérant la présence d'un facteur de susceptibilité dans une région génétique donnée<sup>171</sup>.

Ce type d'étude est souvent critiquée car la plupart des associations mises en évidence n'ont pu être répliquées dans des populations indépendantes<sup>172</sup>. Cette approche exclut les gènes dont le rôle est inconnu ou dont l'implication dans la maladie est insoupçonnable, ce qui constitue une limite supplémentaire à leur utilisation. Pour remédier à ces limites, des approches génome-entier ont été développées.

### II.3.3.2 Approche génome entier

Les progrès importants réalisés dans les techniques de génotypage au cours des années 2000 ont rendu possible de génotyper pour un faible coût plusieurs centaines de milliers de polymorphismes, grâce à la technique de « puces à ADN »<sup>173</sup>. Ces progrès ont permis de s'affranchir de la nécessité de poser une hypothèse biologique préalable, ouvrant la voie aux études d'association génome-entier (Genome Wide Association Study ou GWAS). Ces GWAS testent l'association entre un très grand nombre de marqueurs (pouvant dépasser le million) répartis sur tout le génome et une pathologie donnée, le plus souvent dans de grandes cohortes de patients et de témoins. Depuis les premières associations identifiées en 2005 dans la dégénérescence maculaire liée à l'âge<sup>174</sup>, plus de 13 000 locus de susceptibilité à des maladies ou à des traits phénotypiques ont été recensés<sup>175</sup>.

Ce type d'approche présente toutefois plusieurs limites. Tout d'abord, le fait de tester un très grand nombre de marqueurs nécessite d'adapter le niveau de significativité. Le seuil communément admis pour les GWAS est une valeur  $p$  inférieure à  $5 \times 10^{-8}$ , ce qui implique donc de recruter un grand nombre d'individus pour atteindre une puissance statistique suffisante<sup>176</sup>. Une autre faiblesse de ce type d'étude est de tester l'implication de variants relativement fréquents dans la population, ce qui peut expliquer que la plupart des variants identifiés n'exercent généralement qu'une influence modeste sur la susceptibilité globale à la maladie et ne rendent compte que d'une faible fraction de l'héritabilité totale<sup>177</sup>.

Il apparaît donc nécessaire d'explorer d'autres pistes afin d'identifier cette « hérabilité manquante ». Les GWAS ont jusqu'à présent essentiellement testé des variants communs<sup>178</sup>. Toutefois, l'implication de variants rares dans les maladies complexes semble non négligeable<sup>179</sup>. Le développement récent des techniques de séquençage nouvelle-génération a ouvert la voie à l'étude de ces variants rares. Ainsi, un reséquençage haut-débit des gènes dans lesquels des polymorphismes avaient été associés aux MICI a permis d'identifier près de 70 variants rares associés à la maladie avec un impact fonctionnel potentiel<sup>180</sup>. Les techniques de séquençage de l'ensemble des régions codantes du génome (whole-exome sequencing), voire même de l'ensemble du génome (whole-genome sequencing), sont prometteuses, même si les méthodes d'analyse les plus efficaces restent encore à préciser<sup>181</sup>.

### II.3.4 Contrôle-qualité des données

Les études d'association génèrent souvent de grande quantité de données de génotypage, qui doivent être soumises à un contrôle-qualité rigoureux avant toute analyse. Le contrôle-qualité des données d'une étude d'association (et notamment d'un GWAS) consiste à vérifier trois points principaux : la qualité des échantillons d'ADN, la qualité des marqueurs génotypés et l'absence de biais de sélection dans la population étudiée.

#### II.3.4.1 Évaluation de la qualité des échantillons d'ADN

Dans un premier temps, il est nécessaire de chercher à détecter des problèmes d'identité des échantillons, par vérification du sexe si des données de génotypage du chromosome X sont disponibles ou par vérification des liens de parenté en cas d'étude familiale<sup>182,183</sup>. Une fois ces premières vérifications effectuées, les échantillons ayant un taux de génotypage trop faible sont éliminés de l'analyse<sup>184</sup>.

#### II.3.4.2 Évaluation de la qualité des génotypes

Ensuite, il est nécessaire d'évaluer le taux de génotypage de chaque marqueur et d'éliminer ceux n'atteignant pas un certain seuil fixé le plus souvent entre 95 et 99%<sup>184</sup>. La qualité du génotypage peut aussi être vérifiée par la duplication de certains échantillons d'ADN permettant d'éliminer les SNPs dont la reproductibilité est insuffisante.

L'équilibre de HW doit également être testé pour chaque marqueur. La loi de HW postule qu'au sein d'une population idéale (c'est-à-dire de grande taille, où les unions se font au hasard et en l'absence de migration, de mutation ou de sélection), l'équilibre des fréquences alléliques et génotypiques est conservé d'une génération à l'autre et que les fréquences génotypiques peuvent être déduites des fréquences alléliques (**Figure 12**)<sup>185,186</sup>. Une déviation significative de cet équilibre

peut indiquer la présence d'erreurs de génotypage<sup>187</sup> ou une stratification de population<sup>188</sup>. Un déséquilibre de HW témoigne parfois également d'une réelle association<sup>189</sup> : pour cette raison, de nombreux auteurs proposent de ne tester cet équilibre que chez les sujets sains et d'éliminer les marqueurs qui sont en déséquilibre de HW ( $p < 0,0001$  le plus souvent) dans cette population.

Pour un marqueur avec 2 allèles A et a ayant pour fréquences respectives p et q, on obtient la loi de distribution suivante :

$$p^2 + q^2 + 2pq = 1$$

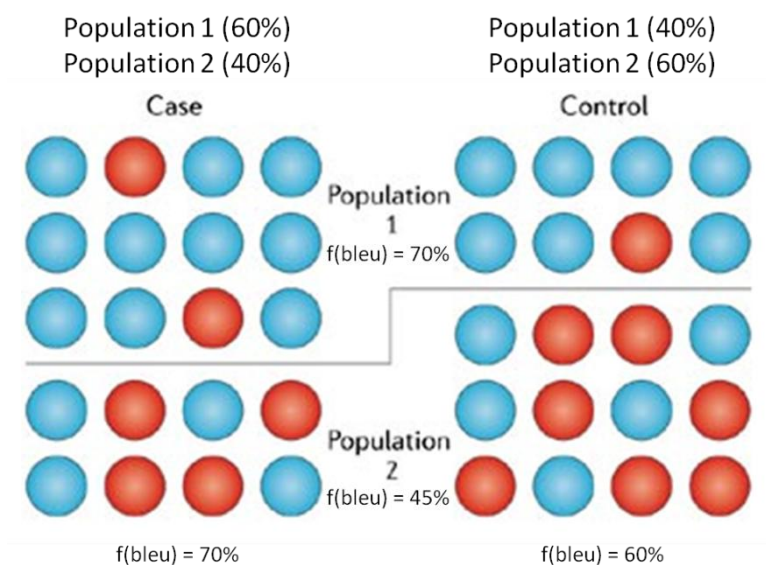
Il est alors possible de vérifier que la distribution génotypique d'une population est conforme à cette loi en la comparant à la distribution théorique grâce à un test du  $\chi^2$ .

**Figure 12 – Équilibre de Hardy-Weinberg**

#### *II.3.4.3 Recherche d'un biais de stratification*

Après avoir vérifié l'intégrité et la qualité de chaque échantillon séparément, il est important de les étudier dans leur ensemble afin de détecter un éventuel biais de stratification. Une population est dite stratifiée s'il existe une différence systématique des fréquences alléliques entre les cas et les témoins en rapport avec leur origine ethnique plutôt qu'avec leur trait phénotypique (**Figure 13**)<sup>190</sup>. Le fait de ne pas détecter une telle stratification peut conduire à l'identification de fausses associations liées aux différences ethniques plutôt qu'au trait étudié<sup>191</sup>.

Plusieurs méthodes statistiques, comme le contrôle génomique, la structure d'association ou l'analyse en composante principale, ont été développées pour détecter la stratification et ajuster les analyses d'association en conséquence<sup>192-194</sup>. Un autre moyen de s'affranchir du biais de stratification est d'utiliser une approche intra-familiale.



**Figure 13 – Biais de stratification ; d'après Balding *et al.*<sup>167</sup>**

	Cas	Contrôles
Allèle A	A	B
Allèle a	C	D

$$OR_{AllèleA} = \frac{\text{probabilité d'être malade avec l'allèle A}}{\text{probabilité d'être malade avec l'allèle a}} = \frac{(A/B)}{(C/D)} = \frac{A * D}{B * C}$$

**Figure 14 – Calcul de l'odds ratio**

### II.3.5 Interprétation des résultats

L'interprétation des résultats est elle aussi soumise à des règles strictes afin de limiter le risque de conclusions erronées.

#### II.3.5.1 Significativité et force de l'association

L'interprétation des résultats d'une analyse d'association nécessite de bien distinguer deux concepts différents : la significativité de l'association et la force de celle-ci.

La significativité de l'association est estimée par la valeur p, ou niveau de probabilité : elle correspond à la probabilité d'obtenir la différence observée entre les deux groupes du simple fait du hasard. Cette valeur est comprise entre 0 et 1 : plus elle se rapproche de 0, plus la significativité est importante.

La significativité n'est toutefois pas un reflet de la force de l'association. Celle-ci est évaluée par l'odds ratio (OR), ou rapport de cote (**Figure 14**). Lorsque l'OR est inférieur à 1, le variant est associé à une diminution du risque de la maladie, tandis que s'il est supérieur à 1, il est associé à une augmentation du risque. L'OR est souvent donné avec son intervalle de confiance, reflétant le degré de certitude d'une estimation.

### II.3.5.2 Correction pour tests multiples

L'interprétation de la p-value doit également tenir compte du nombre de marqueurs testés. La plupart des études d'association testent un grand nombre de marqueurs. Si le seuil de risque de type I alpha (c'est-à-dire le risque de rejeter à tort l'hypothèse nulle) n'est pas adapté au nombre de tests effectués, le risque de faux-positifs augmente donc<sup>195</sup>. De nombreuses méthodes statistiques ont été développées pour prendre en compte la problématique des tests multiples. Il est possible de les séparer schématiquement en deux groupes : « Family Wise Error Rate », ou FWER, et « False Discovery Rate », ou FDR.

Les méthodes FWER contrôlent la probabilité de n'avoir aucun faux-positif et sont donc très conservatrices. La plus fréquemment appliquée est la correction de Bonferroni qui consiste à fixer le seuil de significativité alpha en divisant le seuil  $\alpha$  choisi par le nombre  $n$  de tests réalisés<sup>196,197</sup>. Les méthodes de type FDR permettent quant à elles de contrôler le taux de faux-positifs et sont donc moins conservatrices. Les deux procédures FDR les plus utilisées sont celles de Benjamini & Hochberg<sup>198</sup> et de Benjamini & Yekutieli<sup>199</sup>.

Les méthodes d'identification de gènes de susceptibilité aux maladies complexes ne se restreignent pas à l'étude de la séquence génétique. L'étude de l'expression des gènes constitue une autre méthode d'identification des gènes potentiellement impliqués dans la survenue d'une maladie.

## **II.4. Des polymorphismes génétiques à l'expression des gènes**

L'étude du transcriptome dans les maladies complexes s'est dans un premier temps principalement intéressée à l'identification de gènes différentiellement exprimés. Plus récemment, les données génétiques et transcriptomiques ont été combinées entre elles afin de mieux expliquer les conséquences fonctionnelles des polymorphismes de susceptibilité.

### II.4.1 Étude du transcriptome dans les maladies complexes

Le transcriptome est constitué de l'ensemble des ARNm issus de la transcription des gènes. Son étude s'est récemment développée dans les maladies complexes afin d'identifier des gènes

différentiellement exprimés entre malades et témoins. Elle a aussi permis de mieux comprendre l'implication de certains polymorphismes de susceptibilité. En effet, des variations de la séquence nucléotidique peuvent avoir un impact direct sur l'expression de certains gènes et ainsi favoriser une maladie.

#### *II.4.1.1 Définition*

La transcription est une étape intermédiaire indispensable à la production de protéines à partir de l'ADN génomique. L'ADN doit d'abord être transcrit en pré-ARN ou ARN primaire. Une étape d'épissage vient ensuite, consistant en l'excision des régions non codantes (introns) pour aboutir à la formation d'un ARNm. L'ensemble de ces ARNm issus de la transcription des gènes constitue donc le transcriptome. L'étude du transcriptome permet d'établir un profil d'expression génique spécifique à l'échantillon étudié (type cellulaire ou tissu) à un moment et dans une condition physiologique donnés. Elle peut donc apporter des informations précieuses sur certains états pathologiques. Par exemple, la comparaison du transcriptome entre les cellules provenant de sujets malades et celles provenant de sujets sains peut permettre l'identification des gènes différentiellement exprimés en lien avec la physiopathologie de la maladie.

#### *II.4.1.2 Techniques d'étude*

L'évaluation de l'expression des gènes, initialement limitée à l'étude de quelques gènes candidats, a pu ensuite se généraliser à l'ensemble du transcriptome. Les plus anciennes techniques d'étude génome-entier sont basées sur l'hybridation d'ARNm à des puces<sup>200</sup>. Ces méthodes présentent toutefois plusieurs limites. Tout d'abord, seuls les transcrits déjà connus peuvent être identifiés<sup>201</sup>. Ensuite, des hybridations non spécifiques peuvent survenir et occasionner un bruit de fond parasitant le signal recherché<sup>202</sup>. Enfin, l'efficacité de détection des puces est limitée, ce qui ne permet pas de détecter facilement les transcrits de faible abondance<sup>203</sup>.

Plus récemment ont été développées des méthodes de séquençage haut-débit<sup>204</sup>. Ces méthodes permettent la quantification des transcrits avec la possibilité d'en mesurer l'expression spécifique d'allèles<sup>205</sup>. Elles apportent également un grand nombre d'informations supplémentaires par rapport aux méthodes d'hybridation, telles que l'identification des sites d'initiation de la transcription, de transcrits alternatifs, de fusion de gènes ou encore d'ARN non codants<sup>206</sup>.

#### *II.4.1.3 Stratégie d'étude*

Le plan d'expérience et la stratégie d'analyse d'une étude transcriptomique dépendent avant tout de l'objectif fixé. Concernant les maladies complexes, cet objectif consiste le plus souvent en l'identification des gènes différentiellement exprimés entre malades et témoins. Dans un premier



temps sont déterminés le type d'échantillon biologique à étudier, le traitement de ces échantillons (culture, stimulation,...) et leur nombre.

Une fois le plan d'expérience fixé, l'ARN extrait à partir des échantillons collectés est hybridé sur puce ou séquencé. Les données d'expression brutes obtenues sont ensuite pré-traitées, afin de les rendre analysables<sup>207</sup>. Ce pré-traitement comporte la soustraction du bruit de fond du signal obtenu, puis la normalisation des données permettant de comparer les puces entre elles. Enfin, le signal de l'ensemble des sondes correspondant à un gène est combiné pour obtenir le niveau global d'expression d'un gène.

Les données d'expression sont alors prêtes à être analysées. En fonction du plan d'expérience, plusieurs tests statistiques sont susceptibles d'être utilisés comme le t-test, l'ANOVA (Analysis of Variance) ou la régression linéaire<sup>208</sup>. Les résultats doivent être interprétés en fonction de leur niveau de significativité (donné par la p-value) et de l'importance de la différence observée, évaluée par le fold-change (niveau moyen d'expression d'un groupe rapporté à celui du groupe comparateur).

Ces résultats nécessitent ensuite d'être validés. La première étape de validation est dite « technique » et a pour but de s'assurer de la reproductibilité des valeurs d'expression obtenues. Pour cela, le niveau d'expression des gènes identifiés comme différentiellement exprimés est évalué par une méthode plus robuste, la PCR quantitative, à partir des mêmes échantillons d'ARN. Dans un second temps, le niveau d'expression des gènes d'intérêt est mesuré dans de nouveaux échantillons afin de tenter de confirmer l'expression différentielle (validation biologique).

Les données génétiques et transcriptomiques peuvent être analysées simultanément à la recherche de locus modulant le niveau d'expression d'un gène (expression quantitative trait locus ou eQTL).

## II.4.2 Locus associé à l'expression d'un caractère quantitatif

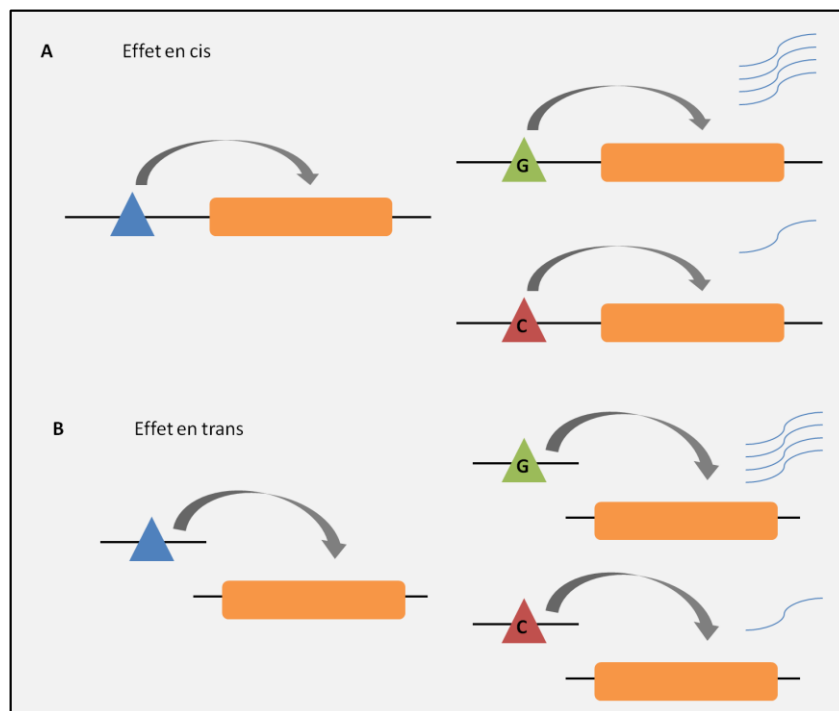
### II.4.2.1 Définition

Les polymorphismes génétiques situés dans des éléments régulateurs ont la capacité potentielle d'influencer le niveau d'expression génique. L'étude simultanée du génome et du transcriptome peut donc permettre une meilleure compréhension des mécanismes à l'origine des maladies complexes.

Un polymorphisme génétique est capable d'influencer le taux d'expression d'un transcrit de deux manières différentes, dites en cis ou en trans (**Figure 15**). Les mécanismes en cis correspondent aux situations où un polymorphisme à proximité du gène (le plus souvent au niveau du site

promoteur) affecte directement la fixation de facteurs de transcription au génome. Dans ce cas, le SNP se situe sur le même chromosome que le gène dont il régule l'expression à une faible distance de celui-ci. À l'inverse, les mécanismes en trans correspondent à une action indirecte du polymorphisme sur l'expression d'un gène via l'intervention d'un élément régulateur (facteur de transcription, co-activateur de la transcription, micro-ARN). Dans ce cas, le SNP responsable de l'association peut se situer n'importe où sur le génome.

Les locus modulant l'expression d'un gène sont appelés eQTLs. Par abus de langage, ce terme est parfois aussi utilisé pour les gènes dont l'expression est modulée par un SNP. On parle alors d'eSNP pour le SNP affectant l'expression.



**Figure 15 – Mécanismes des eQTLs ; d'après Cheung *et al.*<sup>209</sup>**

#### II.4.2.2 Techniques d'étude

L'analyse des eQTLs nécessite de disposer de données d'expression génique dans des cellules, ou dans un tissu cible, chez plusieurs individus. La méthode d'identification des eQTLs est basée sur le principe que les niveaux d'expression génique peuvent être analysés de la même façon que n'importe quel trait phénotypique quantitatif.

L'identification d'un eQTL nécessite de disposer pour un même individu de données génétiques et de données d'expression afin de pouvoir lier les deux. Grâce aux techniques haut-débit, il est maintenant possible de réaliser des études pangénomiques.

#### *II.4.2.3 Stratégie d'analyse*

Tout comme dans les études de transcriptomique, la première étape consiste à choisir le type cellulaire ou tissulaire à analyser. Ce choix est important car il a été démontré qu'une forte proportion des eQTLs étaient dépendants du type cellulaire<sup>210,211</sup>. Bien que ces échantillons soient parfois difficiles à obtenir, le choix d'un type cellulaire ou tissulaire potentiellement impliqué dans la maladie étudiée est donc à privilégier<sup>212</sup>.

Il existe plusieurs méthodes d'identification des eQTLs. La plupart d'entre elles ont été adaptées à partir des méthodes d'analyse utilisées pour la recherche de locus associés à un trait qualitatif : analyses de liaison<sup>213</sup>, analyses d'association chez des individus non apparentés<sup>214</sup> ou analyses d'association intra-familiale<sup>215</sup>. D'autres méthodes ont été développées spécifiquement comme les méthodes spécifiques d'allèle permettant de détecter des différences d'expression entre chacun des allèles chez un individu hétérozygote<sup>216</sup>.

L'identification des facteurs génétiques dans les maladies complexes est rendue aujourd'hui possible grâce au progrès des techniques de biologie moléculaire de haut-débit et à la mise au point de méthodes statistiques adaptées. Toutes ces techniques ont été appliquées à la SpA et ont abouti à l'identification de nombreux facteurs génétiques de susceptibilité.

## CHAPITRE 3 - LA SPONDYLOARTHRITE : UNE MALADIE GÉNÉTIQUE COMPLEXE

---

Les études d'épidémiologie génétique portant sur la SpA ont permis d'établir son caractère multifactoriel et de mieux comprendre l'influence des facteurs génétiques de susceptibilité à la maladie. Un certain nombre de ces facteurs ont pu être identifiés. Plus récemment, les modifications d'expression génique impliquées dans la survenue de la SpA ont fait l'objet de recherches.

### III.1. Épidémiologie génétique de la spondyloarthrite

La plupart des études d'épidémiologie génétique se sont intéressées à la SA, mais plus récemment des recherches ont été menées sur l'ensemble des sous-types de SpA. Si celles-ci ont permis d'établir le poids prédominant des facteurs génétiques dans la survenue de la SpA, certains facteurs environnementaux de susceptibilité ont également été mis en évidence.

#### III.1.1 Épidémiologie génétique de la spondylarthrite ankylosante

La première étape de l'étude génétique d'une maladie complexe repose sur la mise en évidence d'une agrégation familiale de la maladie. Il s'agit ensuite de montrer que cette agrégation est liée à une composante génétique, puis de la caractériser avant d'identifier les gènes en cause (**Figure 16**).

##### III.1.1.1 Agrégation familiale

L'éventualité de l'implication de facteurs génétiques dans la survenue de la SA a été évoquée dans les années 1930, sur la base d'observations de cas regroupés au sein d'une même famille<sup>217,218</sup>. Dans les années 1960, les premières études familiales systématiques ont permis d'évaluer l'agrégation familiale au travers du calcul de l'indice de récurrence  $\lambda$ <sup>219,220</sup>. Cet indice est défini comme le rapport de la prévalence de la maladie chez les individus apparentés à un malade sur la prévalence de la maladie dans la population générale<sup>221</sup> (**Figure 16A**).

Une méta-analyse des données familiales dans la SA publiées avant 2000 a permis d'estimer l'indice de récurrence  $\lambda_1$  chez les apparentés au premier degré (parents/enfants) à 79, similaire à l'indice de récurrence  $\lambda_s$  chez les germains (frères/sœurs) qui était de 82<sup>222</sup>. Cet indice est beaucoup plus élevé que dans la plupart des autres maladies complexes<sup>223</sup>. Il convient toutefois de préciser que la prévalence de SA dans la population générale utilisée dans cette méta-analyse pour estimer

les indices de récurrence étaient plus faibles que dans la plupart des études de prévalence (0,1%), ce qui a pu conduire à une surestimation de l'indice de récurrence.

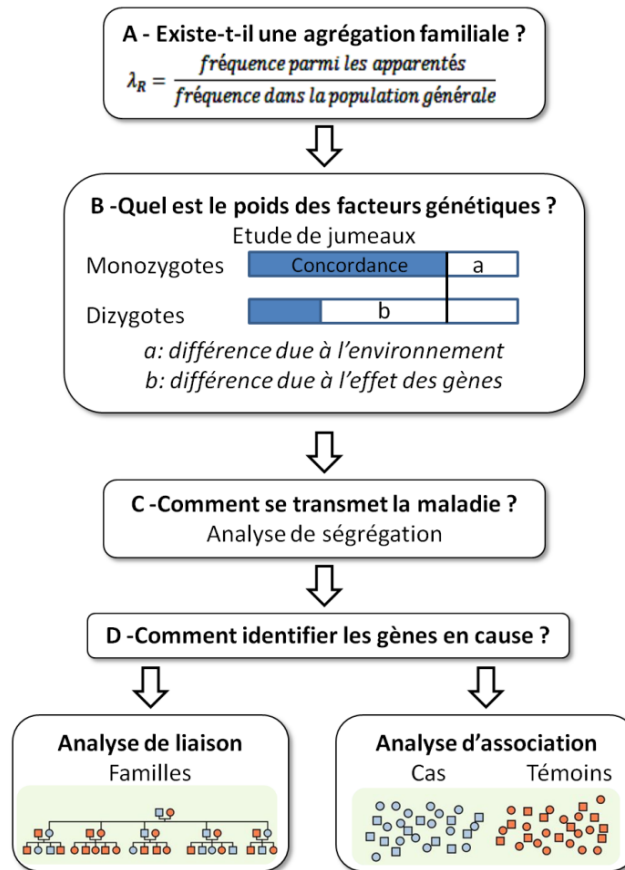


Figure 16 – Les différentes étapes de l'épidémiologie génétique des maladies complexes

### III.1.1.2 Études de jumeaux

L'agrégation familiale peut être le reflet de facteurs environnementaux ou génétiques. L'étude du taux de concordance entre jumeaux permet d'évaluer le poids respectif de ces deux types de facteur. Ce taux correspond au rapport entre le nombre de paires de jumeaux présentant deux individus atteints (jumeaux concordants) et le nombre de paires de jumeaux ne présentant qu'un seul individu atteint (jumeaux discordants). Schématiquement, on considère que la discordance entre jumeaux monozygotes (MZ) est d'origine environnementale, alors que la discordance entre jumeaux dizygotes (DZ) serait d'origine mixte (génétique et environnementale). Ainsi la différence entre les taux de concordances observés chez les jumeaux MZ et DZ est le reflet de l'importance des facteurs génétiques (**Figure 16B**).

Trois études ont évalué le taux de concordance chez des jumeaux atteints de SA<sup>224–226</sup>. Les effectifs de ces études sont globalement faibles mais toutes s'accordent sur un point : le taux de concordance de SA est plus élevé chez les jumeaux monozygotes que chez les jumeaux dizygotes (**Tableau 6**). Cette différence indique un poids important des facteurs génétiques dans la survenue de la maladie par rapport aux facteurs environnementaux (ou l'intervention de facteurs environnementaux ubiquitaires). L'héritabilité de la SA, c'est-à-dire la proportion de la variance expliquée par des effets génétiques, a ainsi été estimée à plus de 90%<sup>225</sup>.

**Tableau 6 – Études de jumeaux réalisées dans la SA**

Étude	Concordance jumeaux MZ (%)	Concordance jumeaux DZ (%)
Järvinen <i>et al.</i> <sup>224</sup>	3/6 (50%)	3/20 (15%)
Brown <i>et al.</i> <sup>225</sup>	6/8 (75%)	4/32 (12.5%)
Pedersen <i>et al.</i> <sup>226</sup>	2/5 (40%)	1/23 (4%)

### III.1.1.3 Études de ségrégation : modélisation de la transmission de la maladie

À partir de la méta-analyse des données familiales précédemment citée, Brown *et al.* ont proposé un modèle de transmission de la SA<sup>222</sup>. En se basant sur les indices de récurrence estimés chez différents types d'apparentés et sur les travaux de Risch concernant la modélisation de la transmission des modèles multi-locus<sup>221</sup>, le modèle de transmission correspondant le mieux à la SA s'est révélé être un modèle multiplicatif oligogénique à cinq locus avec un effet polygénique résiduel. Ce modèle a toutefois tendance à être remis en question : en effet, plus d'une vingtaine de locus de susceptibilité à la SA ont déjà été identifiés et l'ensemble de ces locus est loin d'expliquer la totalité de l'héritabilité à la maladie<sup>227</sup>.

Les études décrites ci-dessus n'ont inclus que des cas de SA. Cependant, il existe d'autres sous-types de SpA parmi les apparentés de patients atteints de SA. Pour avoir une vision exhaustive du terrain génétique de la SpA, il convient de prendre en compte ces cas.

### III.1.2 Épidémiologie génétique de la spondyloarthrite

La majorité des études d'épidémiologie génétique portant sur la SpA dans son ensemble a été réalisée par le Groupe Français d'Étude Génétique de la Spondylarthrite (GFEGS). Après avoir établi la pertinence de l'étude groupée des différents sous-types de SpA, le GFEGS a évalué l'agrégation familiale dans la SpA.

### III.1.2.1 Pertinence de l'étude de la spondyloarthrite dans son ensemble sur le plan génétique

Afin de savoir si l'ensemble des sous-types de SpA pouvait être regroupé dans les études génétiques, le GFEGS a conduit plusieurs études phénotypiques de familles multiplex de SpA, c'est-à-dire possédant plus d'un membre malade<sup>228-230</sup>. Les principales conclusions de ces études sont que les sous-types de SpA sont distribués aléatoirement au sein des familles et que toutes les manifestations de SpA apparaissent liées entre elles<sup>228,229</sup>. Il est donc possible de considérer les sous-types de SpA comme des variations phénotypiques d'une seule maladie, justifiant leur regroupement dans les études génétiques.

L'étude des familles multiplex par la méthode d'analyse en cluster a aussi permis d'individualiser deux formes de SpA. Le phénotype A se caractérisait par un âge de début tardif et la prédominance de l'atteinte axiale, tandis que dans le phénotype B, l'âge de début était plus précoce et la fréquence des atteintes articulaires périphériques et extra-articulaires était supérieure. Ces deux phénotypes, qui pourraient correspondre à différents degrés de sévérité de la maladie, présentaient une tendance à l'agrégation familiale<sup>230</sup>.

### III.1.2.2 Agrégation familiale dans la spondyloarthrite

Contrairement à la SA, une seule étude familiale systématique de la récurrence considérant la SpA dans son ensemble a été réalisée à ce jour<sup>231</sup>. Le taux de récurrence de la maladie était d'environ 12% chez les apparentés au premier degré et chez les germains. En prenant en compte la prévalence estimée de SpA en France en 2001 de 0,3%<sup>2</sup>, les indices de récurrence  $\lambda_1$  et  $\lambda_s$  étaient donc de 40.

Les études de jumeaux décrites précédemment ont montré une concordance incomplète chez les jumeaux monozygotes, ce qui suggère le rôle de facteurs environnementaux dans le déterminisme de la SpA. Peu d'entre eux ont cependant été identifiés.

### III.1.3 Facteurs environnementaux

Les micro-organismes figurent parmi les facteurs environnementaux les plus souvent incriminés dans la survenue de la SpA. Ce rôle est suggéré par le déclenchement d'arthrites réactionnelles faisant suite à des infections uro-génitales ou digestives<sup>91</sup>. L'implication du HLA-B27 est renforcée par le modèle animal du rat transgénique pour le HLA-B27 et la  $\beta$ 2-microglobuline humaine<sup>232</sup>. En effet, ces animaux ne développent aucun symptôme s'ils sont élevés en atmosphère stérile<sup>233</sup>, alors que la reconstitution de leur flore intestinale suffit à la survenue de la maladie<sup>234</sup>.

D'autres facteurs environnementaux semblent intervenir, si ce n'est dans la survenue de la SpA, du moins dans son activité ou sa sévérité. Ainsi, le tabagisme a été récemment suspecté d'être à l'origine d'un début plus précoce de la maladie, d'un niveau d'activité plus élevé et d'une plus grande fréquence de lésions structurales<sup>235,236</sup>. De même, le stress semble augmenter le niveau d'activité de la SpA<sup>237</sup>.

Devant la constatation de la forte héritabilité de la SpA et de son caractère polygénique, de nombreuses études ont cherché à identifier les gènes de susceptibilité à la maladie avec un succès croissant ces dernières années.

### **III.2. Facteurs de susceptibilité génétique identifiés dans la spondyloarthrite**

Le premier facteur génétique de susceptibilité identifié dans la SpA s'est révélé être l'antigène HLA-B27 situé dans le complexe majeur d'histocompatibilité (CMH). Durant les années 1990, plusieurs études de liaison ont été réalisées sans toutefois obtenir de résultats déterminants. Il a fallu attendre 2007 et la publication du premier GWAS pour que des gènes de susceptibilité à la SpA soient identifiés en dehors du CMH.

#### **III.2.1 Complexe majeur d'histocompatibilité**

Le CMH est une région située sur le chromosome 6 (6p21.3). Elle est formée de trois sous-régions : Human Leucocyte antigen (HLA) de classe I, HLA de classe II et HLA de classe III, toutes très polymorphes. Cette région est caractérisée par une grande richesse en gènes, dont la majorité sont impliqués dans la réponse immunitaire. Le CMH a donc été étudié dans de nombreuses maladies auto-immunes et inflammatoires, conduisant à l'identification de facteurs génétiques de susceptibilité à celles-ci<sup>238</sup>. Dans la SpA, l'antigène HLA-B27 qui appartient au CMH de classe I est le facteur de susceptibilité majeur à la maladie, mais d'autres facteurs situés dans le CMH semblent également impliqués.

##### **III.2.1.1 HLA-B27**

##### ***Association à la spondyloarthrite***

L'association entre l'allèle HLA-B27 et la SpA a été rapportée pour la première fois en 1973. Deux études publiées quasi-simultanément mettaient en évidence un taux élevé (96%) de porteurs de l'allèle HLA-B27 parmi les patients atteints de SA<sup>98,99</sup>. Depuis cette description initiale, l'association entre l'antigène HLA-B27 et les différentes formes de SpA a été confirmée à plusieurs reprises<sup>100</sup>. Dans l'étude de la cohorte GAZEL (cf **Annexe 1**), 75% des patients et 6,9% des témoins



étaient porteurs de l'allèle B27. Le risque relatif de développer une SpA chez les sujets HLA-B27 était de 39<sup>4</sup>.

L'important polymorphisme des molécules du CMH, et particulièrement du HLA-B27, a conduit à l'individualisation de plusieurs centaines de sous-types de l'antigène HLA-B27 (<http://www.ebi.ac.uk/imgt/hla/>). L'allèle ancestral, dont dériveraient les autres par mutation, serait HLA-B\*2705, majoritaire dans les populations caucasiennes<sup>239</sup>. La plupart des sous-types communs, comme B\*2702, B\*2704, B\*2705 et B\*2707, ont été associés à la maladie<sup>240</sup>. Deux sous-types semblent en revanche peu ou pas associés à la maladie : B\*2706 (fréquent en Asie du Sud Est) et B\*2709 (décrit uniquement en Sardaigne et en Italie du Sud)<sup>241,242</sup>.

### *Hypothèses physiopathologiques liant HLA-B27 à la spondyloarthrite*

La principale fonction des molécules du CMH de classe I est de présenter des peptides intracellulaires, dérivés de molécules du soi, de virus ou de bactéries, aux lymphocytes T CD8+ cytotoxiques, participant ainsi à la réponse immune adaptative<sup>243</sup>. Le CMH de classe I joue également un rôle dans la réponse immune innée par sa capacité à interagir avec les récepteurs inhibiteurs KIR (Killer Immunoglobulin Receptor) exprimés à la surface des cellules Natural Killer (NK)<sup>244</sup>.

Bien que l'association entre la SpA et HLA-B27 ait été découverte il y a plus de 40 ans, les mécanismes à l'origine de cette association restent à ce jour largement méconnus. Plusieurs théories ont été proposées sans qu'aucune ne soit définitivement confirmée. Elles sont résumées dans le **Tableau 7**.

#### III.2.1.2 Autres gènes du complexe majeur d'histocompatibilité

D'autres allèles HLA de classe I ont été incriminés dans la susceptibilité à la SpA comme HLA-B\*1403, HLA-B\*60, HLA-B\*39, HLA-B\*38, HLA-B\*40 et HLA-B\*52<sup>239</sup>. Cependant, la plupart de ces études portent sur de faibles effectifs et leurs résultats sont parfois discordants. Il est donc difficile d'en tirer des conclusions définitives. Les allèles pour lesquels les données disponibles sont les plus convaincantes sont HLA-B\*60, qui augmente la susceptibilité à la SpA principalement chez les sujets HLA-B27 positifs (HLA-B27+)<sup>245</sup>, et HLA-B\*1403, qui est associé à la maladie dans les populations subsahariennes où le HLA-B27 est extrêmement rare<sup>239</sup>.

D'autres études ont été consacrées aux gènes du CMH en dehors du CMH de classe I. Leur interprétation est toutefois très délicate, du fait de la présence d'un DL considérable dans la région du CMH et du poids du facteur HLA-B27. La plupart des associations décrites sont vraisemblablement la conséquence d'un DL entre les allèles identifiés et le HLA-B27. Certaines études récentes ont toutefois été réalisées avec une méthodologie permettant de s'affranchir au maximum de l'influence

de l'allèle HLA-B27. En génotypant un grand nombre de SNPs du CMH chez des patients B27+ et des témoins B27+, Diaz-Pena *et al.* ont ainsi pu montrer un rôle probable de HLA-DPA1 dans la susceptibilité à la SA<sup>246,247</sup>. Zhou *et al.* ont quant à eux utilisé une stratégie d'analyse stratifiée sur la présence de HLA-B27 qui a permis de mettre en évidence une influence d'un polymorphisme du gène *MICA* sur le risque de SpA<sup>248</sup>.

L'antigène HLA-B27 et l'ensemble du CMH contribuent donc fortement au déterminisme génétique de la SpA. Toutefois, il existe de nombreux arguments en faveur de l'existence d'autres facteurs génétiques.

**Tableau 7 – Principales hypothèses liant HLA-B27 à la SpA**

Hypothèses	Mécanismes
<i>Théories spécifiques de l'antigène</i>	
Peptide arthritogène	Les molécules HLA-B27 associées à la SpA lieraient spécifiquement des peptides à tropisme articulaire, reconnus par des lymphocytes T auto-réactifs. Des peptides exogènes bactériens, ayant une homologie de séquence avec ces peptides arthritogènes, pourraient avoir initialement activé ces clones auto-réactifs <sup>249</sup> .
Mimétisme moléculaire	Des anticorps dirigés contre des motifs peptidiques bactériens pourraient avoir une réaction croisée avec l'antigène tissulaire HLA-B27. Les bactéries en cause dans le déclenchement des arthrites réactionnelles pourraient ainsi réagir de façon croisée avec l'antigène tissulaire HLA-B27 <sup>250</sup> .
<i>Théories indépendantes de l'antigène</i>	
Anomalies du repliement	La chaîne lourde du HLA-B27 se replie plus lentement que d'autres molécules HLA. L'accumulation de protéines mal repliées dans le réticulum endoplasmique pourrait conduire à un stress cellulaire UPR (Unfolded Protein Response) responsable entre autres de l'activation du facteur de transcription NFκB et de la production de cytokines pro-inflammatoires <sup>251</sup> .
Formation d'homodimères	Les chaînes lourdes du HLA-B27 ont tendance à former des homodimères pouvant être exportés à la surface cellulaire. Ces homodimères pourraient être reconnus par des récepteurs KIR sur les cellules NK et les LT. Cette reconnaissance protégerait les cellules NK de l'apoptose et favoriserait la prolifération et la production d'IL17 par les cellules T <sup>252,253</sup> .

### III.2.1.3 Arguments pour l'implication d'autres facteurs génétiques en dehors du complexe majeur d'histocompatibilité

L'allèle HLA-B27 n'est pas nécessaire à la survenue de la maladie : selon les cohortes étudiées, on note que 5 à 25% des sujets atteints de SpA ne sont pas porteurs de l'allèle HLA-B27<sup>4,240</sup>. En outre, la présence du HLA-B27 n'implique pas systématiquement la survenue de la maladie, loin s'en faut. On estime ainsi que seuls 1 à 10% des porteurs de l'allèle HLA-B27 dans la population générale développent une SpA<sup>4,10,23</sup>. Ce risque est environ 15 fois plus élevé chez les sujets HLA-B27 positifs apparentés au premier degré à un individu atteint de SpA<sup>23</sup>. Ce constat suggère l'intervention d'autres facteurs génétiques.

L'étude de jumeaux apporte un argument supplémentaire, puisque la différence du taux de concordance entre jumeaux MZ et DZ persiste même si on ne considère que les paires de jumeaux HLA-B27<sup>224-226</sup>. Enfin, la sévérité de la maladie semble elle-aussi influencée par des facteurs génétiques autres que HLA-B27. En effet, le BASDAI, le BASFI et le BASRI (respectivement indices d'activité, de sévérité fonctionnelle et de sévérité radiologique de la maladie) ont tous été associés à une forte héritabilité<sup>254,255</sup>. Ces études ayant été réalisées uniquement chez des patients HLA-B27+, l'héritabilité observée est donc attribuable à d'autres facteurs qu'à cet antigène HLA.

Au total, la part attribuable au HLA-B27 est évaluée entre 20 et 50% du déterminisme génétique global<sup>231,256</sup>.

### III.2.2 Locus de susceptibilité à la spondyloarthrite en dehors du complexe majeur d'histocompatibilité

Les méthodes de recherche de gènes de susceptibilité dans les maladies complexes décrites précédemment ont toutes été appliquées à la SpA, qu'il s'agisse des études de liaison ou des études d'association.

#### III.2.2.1 Études de liaison

Trois études de liaison pangénomique ont été précédemment publiées dans la SpA : deux se sont intéressées uniquement à la SA<sup>257-259</sup> et la troisième à l'ensemble des sous-types de SpA<sup>260</sup>. Leurs résultats sont résumés dans le **Tableau 8**. Ces trois études ont fait secondairement l'objet de deux méta-analyses<sup>261,262</sup>.

#### *Criblages génomiques de l'équipe d'Oxford*

La première étude a été réalisée par le groupe d'étude génétique de la SA d'Oxford. Elle a comporté deux étapes successives analysées d'abord séparément puis conjointement<sup>257,258</sup>. L'analyse

conjointe portait sur 185 familles comprenant 255 paires de germains atteints. Le génotypage de marqueurs microsatellites répartis sur tout le génome a été effectué chez 818 individus (445 malades et 373 sujets sains). Le nombre de microsatellites différait en fonction du jeu de familles analysé : 505 marqueurs lors de la première étape (99 familles ; 130 paires de germains affectés), 367 marqueurs lors de la seconde étape (86 familles ; 125 paires de germains atteints). L'analyse du chromosome X effectuée séparément a porté sur 234 des 255 paires de germains atteints<sup>263</sup>.

Les données ont ensuite été analysées selon une méthode non paramétrique (bi-points et multipoints). Les seuils de significativité et de suggestivité du LOD score ont été fixés respectivement à 3,6 et 2,2 en accord avec les recommandations de Lander et Kruglyak<sup>133</sup>. L'analyse multipoints a permis de mettre en évidence deux régions significativement liées à la SA : la région du CMH avec un LOD score de 15,6 et la région 16q23.3 avec un LOD score maximal de 4,7. Par ailleurs, cinq locus supplémentaires ont atteint le seuil de suggestivité dans au moins l'un des deux types d'analyse (bi-points ou multipoints) : 1p, 2q, 9q, 10q et 19q. Aucune liaison n'a été observée au niveau du chromosome X.

### *Criblage génomique du consortium Nord-Américain*

En 2004, le North American Spondyloarthritis Consortium (NASC) a publié une étude de liaison portant sur 180 familles (244 paires de germains atteints)<sup>259</sup>. Comme dans l'étude du groupe d'Oxford, seuls des cas de SA étaient inclus. Un jeu de 400 marqueurs microsatellites a été utilisé initialement pour génotyper les individus. Dans un second temps, 16 nouveaux marqueurs ont été génotypés pour affiner les premiers résultats.

La région du CMH était la seule à obtenir un score de liaison significatif. Aucune région n'atteignait le seuil de suggestivité, même si deux régions sur les bras longs des chromosomes 6 et 11 s'en approchaient.

### *Criblage du Groupe Français d'Étude Génétique de la Spondylarthrite*

Contrairement aux études précédentes, le criblage réalisé par le GFECS s'est intéressé à l'ensemble des sous-types de SpA<sup>260</sup>. Sa seconde particularité est d'avoir inclus des familles plus étendues présentant notamment plusieurs générations de malades. L'étude s'est déroulée en deux étapes. Dans un premier temps, 369 marqueurs microsatellites ont été génotypés dans un premier panel de 65 familles multiplex. Une étude d'extension a ensuite été réalisée en génotypant 21 microsatellites localisés sur les régions d'intérêt identifiées lors de la première étape dans les mêmes familles, ainsi que dans 55 familles multiplex supplémentaires.

Lors de la première étape de l'étude, aucune région n'a atteint le seuil de significativité ou de suggestivité en dehors du CMH (NPL score de 5,29). Cependant, quatre régions d'intérêt potentiel ont été identifiées sur les bras longs des chromosomes 5, 9, 13 et 17. Ces régions ont été examinées plus densément dans la deuxième partie de l'étude qui a conduit à l'identification d'une région significativement liée en 9q31-34 (NPL score de 4,87). Cette région chevauchait la région où une liaison suggestive avait été mise en évidence par le groupe d'Oxford. Ce locus a été baptisé SPA2.

### *Méta-analyses*

Deux méta-analyses ont été publiées successivement en 2005 et en 2007<sup>261,262</sup>. Les deux se basaient sur les résultats des trois études précédemment citées, mais seule l'étude de Carter *et al.*<sup>262</sup> a utilisé les données brutes de génotypage.

Bien que la méthodologie utilisée ait été différente, les résultats des deux études se sont révélés assez proches. La méta-analyse de Lee *et al.*<sup>261</sup> mettait en évidence une liaison significative au niveau de la région du CMH, ainsi qu'un signal de liaison suggestif dans cinq autres régions : 6q, 9q, 16q, 17p et 19p. L'étude de Carter *et al.* identifiait une liaison significative avec le CMH et une liaison suggestive avec trois autres locus, 6q, 10q et 16q.

**Tableau 8 – Résultats des analyses de liaison dans la SpA**

Étude	Phénotype	Nombre de Familles	Nombre de Microsatellites	Locus significatifs	Locus suggestifs
Oxford <sup>257,258</sup>	SA	185 (99 + 86)	505+367	CMH, 16q23.3	1p, 2q, 9q, 10q et 19q
NASC <sup>259</sup>	SA	180	400 (+16)	CMH	Aucun
GFEGS <sup>260</sup>	SpA	120	369 (+ 21)	CMH 9q31-34	Aucun
Lee <i>et al.</i> <sup>261</sup>	SA	-	-	CMH	6q, 9q, 16q, 17p et 19p
Carter <i>et al.</i> <sup>262</sup>	SA	-	-	CMH	6q, 10q et 16q

### *Étude approfondie du locus SPA2*

Parmi les régions chromosomiques liées de manière significative ou suggestive à la SpA, seule la région SPA2 a fait l'objet d'une étude plus approfondie. L'objectif était d'identifier le ou les facteurs génétiques de susceptibilité à la maladie expliquant le signal de liaison observé à la fois dans le criblage génomique du GFEGS et dans celui du groupe d'Oxford.

Les premiers résultats de l'étude de ce locus ont été publiés en 2009<sup>264</sup>. Grâce à une analyse de liaison densifiée de la région 9q31-34 (28 marqueurs micro-satellites génotypés dans 149 familles

multiplex dont 120 avaient déjà été analysées dans l'étude de liaison pangénomique), la région d'intérêt a pu être réduite à un intervalle de 13 Mb. Une cartographie fine de déséquilibre de liaison de cet intervalle a été établie grâce au génotypage de 1536 tag SNPs dans 136 familles (36 familles ayant obtenu les scores de liaison les plus élevés dans la région et 100 nouvelles familles). Elle a conduit à l'identification d'un SNP significativement associé à la SpA, rs4979459. Le génotypage de ce SNP et d'une trentaine d'autres l'entourant dans 287 familles a permis de montrer une association significative entre un haplotype de six SNPs et la SpA, association qui a été ensuite répliquée dans une population indépendante de cas-témoins.

Parallèlement à ces investigations, les gènes-candidats de la région ont été re-séquencés en deux étapes. Tout d'abord, les régions codantes du gène de la ténascine C, apparaissant comme le meilleur candidat du fait de sa position et de sa fonction (ce gène code pour une glycoprotéine de la matrice extra-cellulaire exprimée dans les enthèses), ont été re-séquencées. Les polymorphismes les plus intéressants ont été ensuite génotypés dans une collection de trios. Aucune association significative n'a pu être mise en évidence<sup>265</sup>. Dans un second temps, les régions codantes et régulatrices de neuf gènes-candidats positionnels supplémentaires ont été re-séquencées (*ZNF618*, *A1L4R1\_HUMAN*, *AMBP*, *KIF12*, *ORM1*, *ORM2*, *C9ORF91*, *ENSESTG000000230601* et *TNFSF8*). Les polymorphismes différenciellement représentés entre les patients et les témoins ont secondairement fait l'objet d'une étude d'extension dans une population de cas-témoins français indépendants. Les polymorphismes dont l'association avec la maladie a été confirmée ont été génotypés dans deux cohortes, portugaise et belge, dans le but de répliquer les associations. Au terme de ces étapes, il a été montré que rs3181357, un SNP intronique rare du gène *TNFSF8*, était significativement associé à la SpA<sup>266</sup>.

La présence de ce SNP rare ne semblait toutefois pas expliquer à elle seule tout le signal de liaison détecté dans la région. Cette région délimitée par l'haplotype de susceptibilité a fait depuis l'objet d'un re-séquençage complet afin d'identifier d'éventuels facteurs de susceptibilité additionnels.

Les résultats des études de liaison dans la SpA ont été dans l'ensemble assez décevants jusqu'à présent, avec un seul locus de susceptibilité en dehors du CMH identifié dans deux études. L'analyse de ce locus, *SPA2*, montre la difficulté de trouver le ou les gène(s) de susceptibilité expliquant la liaison. Pour ces raisons, les études de liaison ont été progressivement délaissées au profit des études d'association.

### III.2.2.2 Études d'association

Les études d'association ont d'abord été restreintes à l'analyse de gènes-candidats puis, grâce aux progrès technologiques, ont pu s'étendre à l'ensemble du génome.

### *Études gène-candidat*

De nombreuses études d'association gène-candidat dans la SpA ont été publiées, concernant le plus souvent le sous-type SA. Leurs résultats sont résumés dans le **Tableau 9**. Toutefois, comme dans la plupart des maladies complexes, la majorité des associations mises en évidence n'ont pu être répliquées de façon indépendante, remettant en question la réalité de ces associations.

### *Études génome-entier*

À ce jour, seul le sous-type SA a fait l'objet de GWAS. Quatre GWAS ont été publiés : trois dans des populations caucasiennes<sup>267–269</sup> et une dans la population chinoise Han<sup>270</sup>. Ces études ont permis l'identification de facteurs de susceptibilité associés à la SA en dehors du CMH. Ces facteurs sont recensés dans le **Tableau 10**.

Plus récemment, l'International Genetics of Ankylosing Spondylitis consortium (IGAS) a publié les résultats d'une étude gènes-candidats à grande échelle. Cette étude a consisté en l'analyse simultanée d'un grand nombre de locus potentiellement impliqués dans les maladies auto-immunes ou inflammatoires, à l'aide d'une puce dédiée (Immunochip®). Cette vaste étude internationale a inclus plus de 10 000 patients atteints de SA et 15 000 témoins. Outre la confirmation de la plupart des facteurs génétiques identifiés lors des précédents GWAS, 13 nouveaux locus de susceptibilité ont été identifiés (**Tableau 10**)<sup>256</sup>.

Les larges études d'association pangénomique ont donc permis l'identification d'un nombre croissant de locus de susceptibilité à la SA. Il faut toutefois souligner que la plupart des facteurs de susceptibilité modifient très peu le risque de développer la maladie. Ainsi, l'ensemble des facteurs identifiés dans l'étude de l'IGAS explique moins d'un quart de l'héritabilité totale (20,1% pour l'allèle HLA-B27 et seulement 4,3% pour l'ensemble des autres facteurs)<sup>256</sup>. Néanmoins, ces découvertes ont permis d'améliorer la compréhension de la SpA.

**Tableau 9 – Principaux résultats des études d'association gène-candidat dans la SpA**

Gène	Localisation	Principaux résultats
<b>TNAP</b>	1p36	Association identifiée dans une population caucasienne <sup>271</sup> Pas d'association dans deux études chinoises <sup>272,273</sup>
<b>FCRL4</b>	1q21	Association identifiée dans une étude chinoise <sup>274</sup>
<b>FCGR2B</b>	1q23	Association identifiée dans une étude chinoise <sup>275</sup>
<b>IL10</b>	1q32-33	Association identifiée dans une étude chinoise <sup>276</sup> Résultats contradictoires dans la population caucasienne <sup>277,278</sup>
<b>IL1</b>	2q14	Association identifiée dans toutes les études <sup>279-288</sup> sauf une étude coréenne <sup>289</sup> Différences dans les gènes du cluster impliqués selon les études
<b>CTLA4</b>	2q33	Association identifiée dans deux études (iranienne <sup>290</sup> et chinoise <sup>291</sup> )
<b>PDCD1</b>	2q37	Association identifiée dans une étude coréenne <sup>292</sup> et deux études chinoises <sup>293,294</sup> Pas d'association dans une étude iranienne <sup>295</sup>
<b>ANKH</b>	5p15	Résultats contradictoires dans la population caucasienne <sup>296-298</sup> Association identifiée dans une étude chinoise <sup>273</sup>
<b>MSX2</b>	5q35	Association identifiée dans une étude japonaise <sup>299</sup>
<b>CA1</b>	8q21	Association identifiée dans une étude chinoise <sup>300</sup>
<b>ASPN</b>	9q22	Association identifiée dans une étude chinoise <sup>301</sup>
<b>TLR4</b>	9q33	Résultats contradictoires <sup>302-307</sup> Pas d'association identifiée dans une méta-analyse <sup>308</sup>
<b>LRP5</b>	11q13	Association identifiée dans une étude chinoise <sup>309</sup>
<b>MMP3</b>	11q22	Pas d'association dans une étude caucasienne <sup>310</sup> Association identifiée dans une étude taïwanaise <sup>311</sup>
<b>JARID1A</b>	12p11	Association identifiée dans deux études (caucasienne <sup>312</sup> et chinoise <sup>313</sup> )
<b>ORAI1</b>	12q24	Association identifiée dans une étude chinoise <sup>314</sup>
<b>IkBa</b>	14q13	Association identifiée dans une étude taïwanaise <sup>315</sup>
<b>MEFV</b>	16p13	Association identifiée dans trois études turques <sup>316-318</sup>
<b>CARD15</b>	16q21	Aucune association démontrée <sup>319-325</sup>
<b>STAT3</b>	17q21	Association identifiée dans trois études (caucasienne <sup>326</sup> et chinoises <sup>327,328</sup> )
<b>ACE</b>	17q23	Association identifiée dans une étude koweïtienne <sup>329</sup> et une étude turque <sup>330</sup>
<b>TGFB1</b>	19q13	Résultats contradictoires <sup>331-333</sup>
<b>CYP2D6</b>	22q13	Association faible identifiée dans deux études <sup>334,335</sup>



**Tableau 10 – Résultats des études d'association génome-entier**

Les locus indiqués en gras correspondent aux locus répliqués, ceux en bleu à des associations suggestives

Étude	Nombre de cas/témoins (SNPs génotypés)	Facteurs génétiques identifiés en dehors du CMH
<b>Burton <i>et al.</i><sup>267</sup></b>	Découverte : 922/1466 (15 333) Réplication : 471/625 (15)	<i>IL23R</i> , <i>ERAP1</i>
<b>Reveille <i>et al.</i><sup>268</sup></b>	Découverte : 2053/5140 (288 662) Réplication : 898/1518 (163)	<b><i>IL23R</i>, <i>ERAP1</i></b> 2p15, 21q22 <i>ANTXR2</i> , <i>IL1R2</i>
<b>Evans <i>et al.</i><sup>269</sup></b>	Découverte : 3023/8779 (2 223 620) Réplication : 2111/4483 (49)	<b><i>IL23R</i>, <i>ERAP1</i>, 2p15, 21q22</b> , <i>KIF21B</i> , <i>RUNX3</i> , <i>IL12B</i> , <i>LTBR-TNFRSF1A</i> , <i>ANTXR2</i> , <i>PTGER4</i> , <i>CARD9</i> , <i>TBKBP1</i>
<b>Lin <i>et al.</i><sup>270</sup></b>	Découverte : 1837/4231 (1 356 350) Réplication : 2100/3496 (30)	<i>EDIL3-HAPLN1</i> , <i>ANO6</i> , <b>2p15</b>
<b>IGAS <i>et al.</i><sup>256</sup></b>	Découverte : 10 619/ 15145 (128 935)	<b><i>RUNX3</i>, <i>IL23R</i>, <i>GPR25-KIF21B</i>, 2p15,</b> <b><i>PTGER4</i>, <i>ERAP1</i>, <i>IL12B</i>, <i>CARD9</i>, <i>LTBR-</i></b> <b><i>TNFRSF1A</i>, <i>NPEPPS</i>, 21q22</b> <i>IL6R</i> , <i>FCGR2A</i> , <i>UBE2E3</i> , <i>GPR35</i> , <i>BACH2</i> , <i>ZMIZ1</i> , <i>NKX2-3</i> , <i>SH2B3</i> , <i>GPR65</i> , <i>IL27-</i> <i>SULT1A1</i> , <i>NOS2</i> , <i>TYK2</i> , <i>ICOSLG</i>

### III.2.3 Contribution des études génétiques à la compréhension de la spondyloarthrite

L'identification de facteurs de susceptibilité à la SpA a permis de mettre en évidence ou de renforcer l'implication de certains mécanismes dans la survenue de la maladie. Elle a également établi la notion de terrain génétique partagé.

#### III.2.3.1 Voies physiopathologiques impliquées

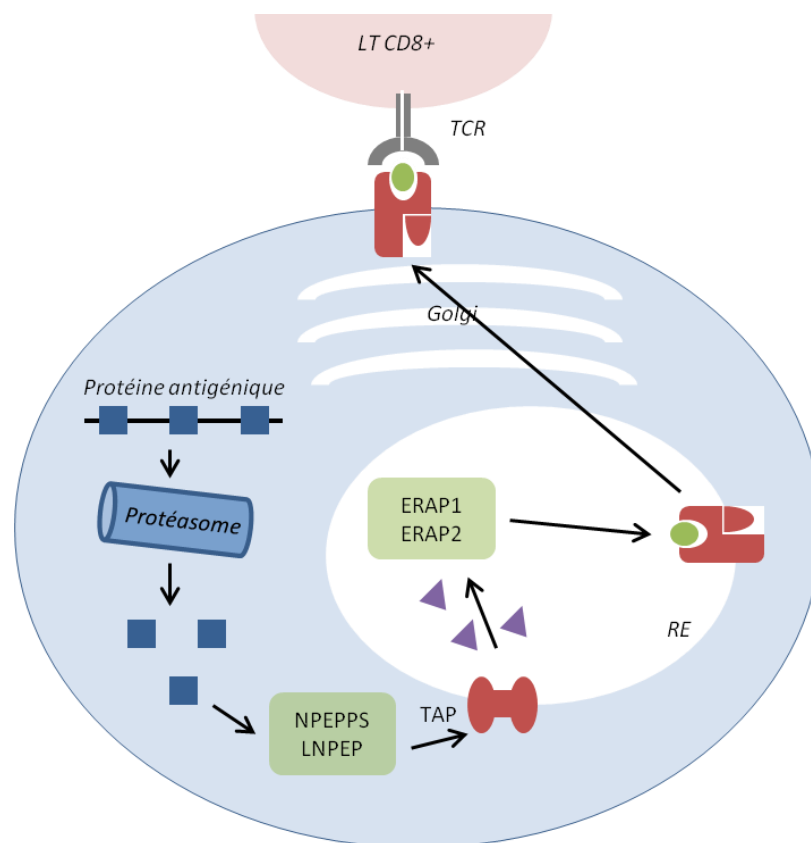
##### Aminopeptidases

Le premier GWAS publié en 2007 a révélé une forte association entre plusieurs polymorphismes d'*ERAP1* et la SA<sup>267</sup>. Cette association a été ensuite répliquée de nombreuses fois, que ce soit dans la population caucasienne ou dans la population asiatique<sup>256,267,268,336–346</sup>. Toutefois, *ERAP1* est situé dans un bloc de DL étendu, incluant plusieurs gènes (*CAST*, *ERAP1*, *ERAP2* et *LNPEP*). Par conséquent, le polymorphisme causal au sein de cette région n'a pu être formellement établi<sup>337</sup>. Le modèle proposé par Evans *et al.* implique rs30187 comme principal responsable de l'association et rs10050860 ou rs17482078 comme responsables d'un effet secondaire<sup>269</sup>.

Le gène *ERAP1* (endoplasmic reticulum aminopeptidase 1), localisé sur le bras long du chromosome 5, code pour une aminopeptidase transmembranaire. Sa fonction principale est le

clivage dans le réticulum endoplasmique des peptides générés par le protéasome afin d'atteindre une longueur optimale pour leur présentation par les molécules du CMH de classe I (**Figure 17**)<sup>347,348</sup>. Une deuxième fonction de clivage de récepteurs des cytokines pro-inflammatoires (TNFR1, IL-1RII, IL-6R) a été décrite<sup>349-351</sup>. Cette fonction a été remise en question, notamment car les souris déficientes pour *erap1* possèdent des niveaux de TNFR1 et d'IL6-R solubles similaires à ceux des souris sauvages<sup>269</sup>.

Ces deux fonctions peuvent aisément être reliées à la SpA. Néanmoins, plusieurs éléments suggèrent que c'est le rôle d'aminopeptidase d'ERAP1 qui est en cause dans la susceptibilité à la SpA. Tout d'abord, une forte interaction génétique a été mise en évidence entre le SNP rs30187 du gène *ERAP1* et HLA-B27<sup>256,269,352,353</sup>. En effet, l'allèle T de ce SNP prédispose à la SpA uniquement chez les porteurs du HLA-B27. Ensuite, d'autres aminopeptidases (*ERAP2*, *LNPEP* et *NPEPPS*) ont été associées à la SpA (**Figure 17**)<sup>256,267,342</sup>. Enfin, Haroon *et al.* ont montré que le clivage des récepteurs des cytokines pro-inflammatoires n'était pas modifié par les polymorphismes de susceptibilité à la SpA<sup>354</sup>.



**Figure 17 – Implication des différentes aminopeptidases associées à la SA ; d'après Parkes *et al.***<sup>355</sup>

Plusieurs études ont cherché à déterminer les conséquences des polymorphismes d'*ERAP1* associés à la SpA sur son activité enzymatique. Elles ont permis de démontrer que les allèles de

susceptibilité de plusieurs polymorphismes non synonymes étaient associés à une augmentation de l'activité enzymatique d'ERAP1<sup>269,356–360</sup>. Cette modulation de l'activité de l'enzyme pourrait être en lien avec des modifications de sa structure tridimensionnelle par ces polymorphismes<sup>357,358</sup>. Les conséquences de l'augmentation de l'activité aminopeptidase d'ERAP1 sont toutefois encore mal connues. Certains auteurs suggèrent que cette hausse est corrélée à une augmentation de l'expression du CMH de classe I à la surface cellulaire<sup>361</sup>, tandis que d'autres privilégient l'hypothèse d'une modification du peptidome du HLA-B27<sup>359,362,363</sup>.

### *Voie de différenciation Th17*

Le second gène de susceptibilité à la SpA identifié en dehors du CMH était celui du récepteur à l'interleukine 23 (*IL23R*)<sup>267</sup>. Cette association a ensuite été répliquée dans de nombreuses populations caucasiennes<sup>339,364–366</sup>. Les polymorphismes associés à la SpA dans les populations caucasiennes ne semblent en revanche pas l'être dans les populations asiatiques<sup>336,367,368</sup>. Des associations avec d'autres polymorphismes de ce gène ont été secondairement mises en évidence dans ces populations<sup>369,370</sup>. Comme pour *ERAP1*, deux polymorphismes semblent expliquer la majorité du signal d'association (rs11209026 et rs11209032)<sup>269</sup>.

Le gène de l'*IL23R*, localisé sur le chromosome 1p31, code pour le récepteur de l'interleukine 23. L'une des fonctions de cette cytokine est de stabiliser et d'amplifier la différenciation des lymphocytes T auxiliaires CD4+ en lymphocytes Th17<sup>371</sup>. Ces lymphocytes Th17 sont caractérisés par la production d'IL17 et d'IL22 et sont associés à une forte réponse pro-inflammatoire<sup>372</sup>. De nombreux arguments suggèrent l'implication de l'axe IL23-Th17 dans la SpA<sup>373</sup>. Parmi ces arguments, on peut citer la constatation d'un taux augmenté d'IL17 dans le sérum des patients atteints de SA<sup>374</sup>, ainsi qu'une augmentation de la fréquence de leurs cellules Th17 circulantes<sup>375</sup>. Des polymorphismes d'autres gènes de l'axe IL23-Th17 (*IL12B*, *CARD9*, *PTGER4*, *TYK2*, *IL6R* et *IL27*) ont également été associés à la SpA<sup>256,269</sup>.

Les conséquences fonctionnelles de certains de ces polymorphismes ont été analysées. Le variant le plus étudié est le SNP non synonyme rs11209026 induisant le remplacement d'une arginine par une glutamine en position 381 (R381Q) à l'origine d'une diminution du risque de SpA. Toutes les études concluent à une perte de fonction induite par l'allèle protecteur de ce polymorphisme. Cette perte de fonction semble affecter à la fois la voie Th17 (avec une diminution de la production d'IL17 et d'IL22 induite par l'IL23, ainsi qu'une diminution de la phosphorylation de STAT3) et la voie Th1 (avec une diminution de l'expression des gènes *IFNG*, *TNFA*, *IL1RB2* et *TBX21*)<sup>376–379</sup>. Par ailleurs l'allèle de susceptibilité de rs4129267, le SNP de l'*IL6R* le plus associé à la SpA dans l'étude de l'IGAS, a été associé à une diminution de la concentration sérique de la forme soluble du récepteur de l'IL6<sup>256,380</sup>.

### III.2.3.2 Spondylarthritis ankylosante, spondyloarthritis et autres maladies inflammatoires : un terrain génétique partagé

La majorité des études génétiques dans la SpA se sont restreintes aux cas de SA. Néanmoins des études plus récentes ont montré que l'association d'un certain nombre de locus à la SA peut être étendue à l'ensemble des SpA. La comparaison avec les études génétiques dans d'autres maladies inflammatoires ou auto-immunes a également permis de faire émerger la notion de terrain génétique partagé.

#### *Généralisation à la SpA des associations mises en évidence dans la SA*

Compte-tenu des résultats des études familiales du GFECS et du partage de certains facteurs génétiques avec d'autres pathologies appartenant au spectre de la SpA, comme la maladie de Crohn ou le psoriasis, il apparaît possible que certains des facteurs associés à la SA soient aussi associés aux autres sous-types de SpA. Toutefois, peu d'études génétiques se sont intéressées à la SpA dans sa globalité<sup>288,353,381-384</sup> et en particulier aux formes non radiographiques de SpA (SpA sans sacro-iliite radiologique selon les critères de New York modifiés<sup>45</sup>).

C'est la démarche qui a été appliquée par notre équipe pour trois gènes de susceptibilité à la SA, le cluster de gènes de l'*IL1*, *ERAP1* et *IL23R*<sup>288,383,384</sup>. Ces études ont montré que certains polymorphismes sont associés à la SpA dans sa globalité, alors que dans d'autres cas, l'association est restreinte à certains sous-types (**Tableau 11**). Ainsi, les polymorphismes d'*ERAP1* testés étaient associés à l'ensemble des SpA<sup>383</sup>, association secondairement confirmée par une étude roumaine<sup>353</sup>. L'étude des polymorphismes au sein du cluster de gènes de l'*IL1* a quant à elle révélé des différences entre sous-types : le SNP rs1894399 du gène *IL1A* semblait plus spécifiquement associé à la SA et le SNP rs3811058 du gène *IL1F10* aux formes non ankylosantes de la maladie<sup>288</sup>. Enfin, le SNP non synonyme de l'*IL23R* rs11209026 est fortement associé à la SA, alors qu'il n'est pas associé aux formes non ankylosantes de la maladie<sup>384</sup>. L'effet sur l'ankylose semble être en lien avec une modulation de l'inflammation systémique, puisque ce SNP est aussi associé au taux de CRP dont l'augmentation est un facteur associé au risque d'ankylose<sup>385</sup>.

Les résultats de ces études tendent donc à montrer qu'il existe un socle de facteurs génétiques communs aux sous-types de SpA, alors que d'autres facteurs ont plutôt un rôle de modulateur phénotypique. Au-delà des différents sous-types de SpA, la notion de terrain génétique partagé avec d'autres pathologies inflammatoires ou auto-immunes commence à émerger.

**Tableau 11 – Résultats des études d'association comparant SA et SpA non radiographiques**

Gène	Effectifs	Polymorphismes	Association (p-value)	
			SA	SpA non SA
<b>Cluster <i>IL1</i><sup>288</sup></b>	185 trios 414 cas/264 contrôles	rs2856836 ( <i>IL1A</i> )	0,024	0,493
		rs17561 ( <i>IL1A</i> )	0,160	0,244
		rs1894399 ( <i>IL1A</i> )	0,026	0,142
		rs3811058 ( <i>IL1F10</i> )	0,055	0,029
<b><i>ERAP1</i><sup>383</sup></b>	France : 414 cas / 384 contrôles Belgique : 317 cas / 248 contrôles	rs17482078	0,035	0,005
		rs10050860	0,001	0,096
		rs30187	3,9x10 <sup>-4</sup>	0,005
<b><i>IL23R</i><sup>384</sup></b>	415 cas / 372 contrôles 383 trios	rs11209026	4,6x10 <sup>-6</sup>	0,41

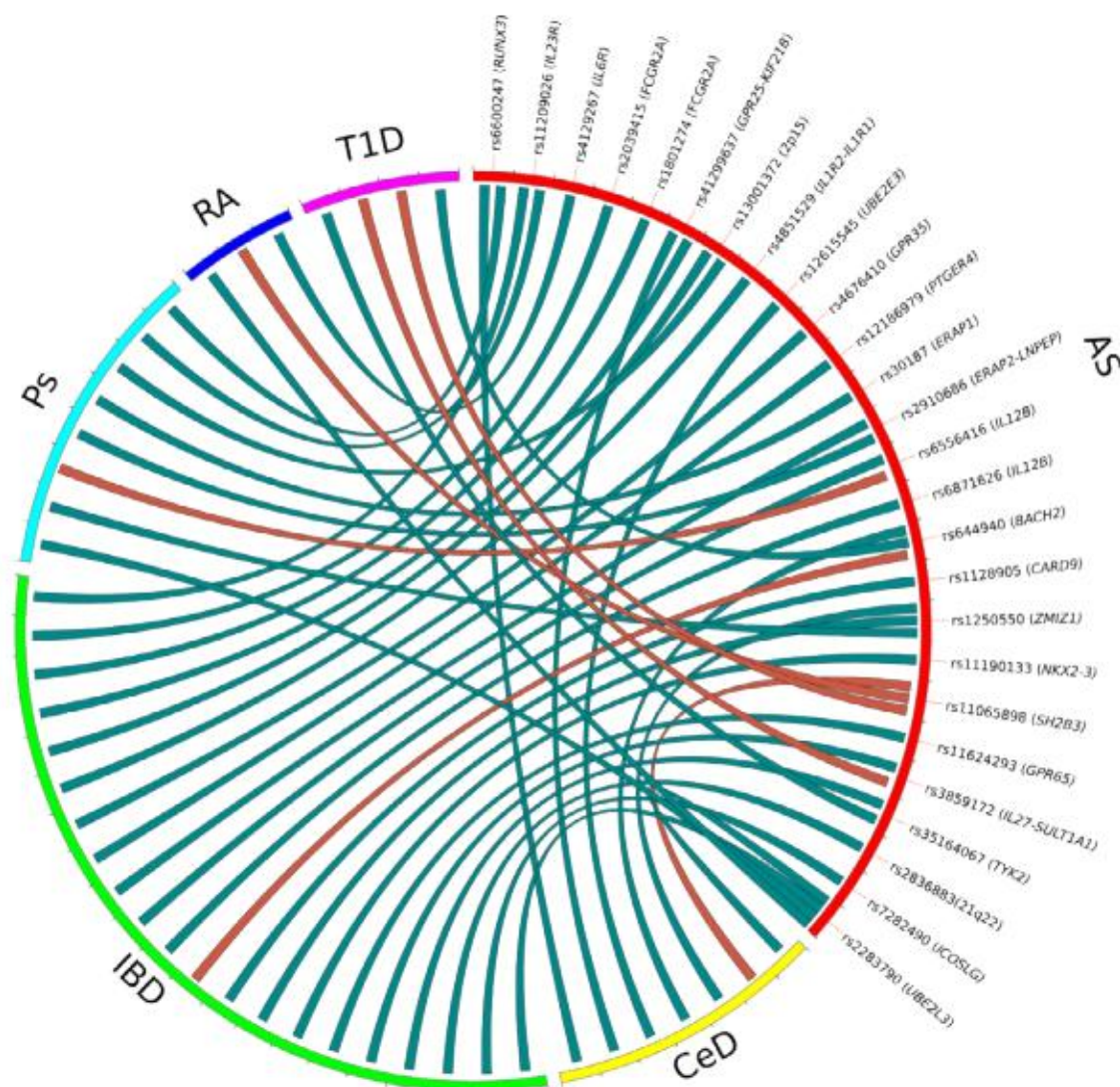
### *Terrain génétique partagé avec d'autres pathologies*

La SpA est fréquemment associée à d'autres pathologies inflammatoires telles que le psoriasis ou les MICI, suggérant l'existence de facteurs génétiques communs entre ces pathologies. Au-delà de ces maladies, d'autres maladies inflammatoires ou auto-immunes semblent également partager des facteurs génétiques de susceptibilité avec la SpA, comme le montre l'analyse comparative de plusieurs études utilisant la puce Immunochip<sup>355</sup>.

Trois types de locus partagés étaient distingués dans cette analyse : les locus corrélés et concordants (même SNP impliqué et même sens de l'association), les locus corrélés mais discordants (même SNP impliqué mais sens contraire de l'association) et les locus non corrélés (même locus mais différents SNPs impliqués). Les résultats comparant la SA aux autres maladies étudiées (maladie cœliaque, MICI, psoriasis, PR et diabète de type 1) sont résumés dans le **Tableau 12** et la **Figure 18**. La SA présente une forte proximité avec les MICI et le psoriasis. De façon plus inattendue, on note également un nombre important de locus partagés entre la SA et la maladie cœliaque, peut-être en lien avec l'inflammation digestive.

**Tableau 12 – Partage des locus de susceptibilité entre la SA et d'autres pathologies**

	Concordants	Discordants	Non corrélés
<b>Maladie cœliaque</b>	6	1	1 ( <i>ICOSLG</i> )
<b>MICI</b>	20	1	2 ( <i>IL23R/IL12RB</i> , <i>TYK2</i> )
<b>Psoriasis</b>	7	1	3 ( <i>IL12B</i> , <i>NOS2</i> , <i>TYK2</i> )
<b>Polyarthrite rhumatoïde</b>	2	1	2 ( <i>FCGR2A</i> , <i>TYK2</i> )
<b>Diabète de type 1</b>	2	2	1 ( <i>GPR25-KIF21B</i> )



**Figure 18 – Partage de locus de susceptibilité ; d’après Parkes *et al.*<sup>355</sup>**

Les liens figurés en bleu correspondent aux locus concordants, ceux en rouge aux locus discordants ; AS : Spondylarthrite Ankylosante ; CeD : maladie coeliaque ; IBD : maladie inflammatoire chronique de l’intestin ; Ps : psoriasis ; RA : polyarthrite rhumatoïde ; T1D : diabète de type 1.

Même si elle a permis une meilleure compréhension de la SpA, l’étude des polymorphismes génétiques est limitée dans les informations fonctionnelles qu’elle peut apporter. Une approche complémentaire permettant de pallier ces insuffisances consiste à étudier l’expression des gènes afin de mieux appréhender les mécanismes en cause dans la SpA.

### **III.3. Étude du transcriptome dans la spondyloarthrite**

Depuis 2002, plus d’une dizaine d’études ont analysé le transcriptome dans la SpA (**Tableau 13**)<sup>386–399</sup>. La grande diversité des stratégies utilisées dans ces études rend difficile leur comparaison

et explique en partie la faible concordance des résultats obtenus. Ces études ont toutefois souligné l'implication potentielle de certaines voies de signalisation dans la physiopathologie de la SpA.

### III.3.1 Différences méthodologiques entre les études

La comparaison des méthodologies utilisées dans les études du transcriptome dans la SpA met en évidence d'importantes différences que ce soit dans leurs objectifs, dans la nature des échantillons étudiés ou dans le choix des témoins utilisés.

#### III.3.1.1 Objectifs de l'étude

La majorité des études transcriptomiques réalisées dans la SpA avaient pour but de mettre en évidence des gènes différentiellement exprimés entre malades et témoins, afin de déterminer les voies de signalisation ou les mécanismes dérégulés dans la pathologie<sup>386–390,392–398</sup>. Leurs principaux résultats sont détaillés ultérieurement.

Dans un but diagnostique, certains auteurs ont cherché à identifier un gène ou un ensemble de gènes permettant de discriminer au mieux les patients atteints de SpA par rapport aux témoins. Ainsi, à partir de l'analyse d'une première série de patients atteints de SA et de témoins sains, Duan *et al.* ont établi une liste de trois gènes (*NR4A2*, *TNFAIP3* et *CD69*) permettant de séparer les patients des témoins. Ils ont ensuite testé les performances de ce jeu de trois gènes dans une deuxième série d'individus avec des résultats satisfaisants (aire sous la courbe ROC (Receiver Operating Characteristic) de 77%)<sup>387</sup>. De façon similaire, le gène *RGS1* semblait être un bio-marqueur diagnostique valide de la SA et des SpA indifférenciées dans l'étude de Gu *et al.*<sup>390</sup>.

L'identification de facteurs associés à une bonne réponse thérapeutique est un autre objectif potentiel des études de transcriptome dans les maladies complexes. Une seule équipe s'est fixé cet objectif dans la SpA. Haroon *et al.* ont étudié le profil transcriptomique d'ARN provenant de sang total de 16 patients atteints de SA avant et deux semaines après l'administration d'Infliximab<sup>391</sup>. Aucun gène n'a permis de distinguer les patients ayant une bonne réponse clinique à l'Infliximab de ceux chez qui ce traitement s'est révélé inefficace.

#### III.3.1.2 Choix du type cellulaire/tissulaire

Le choix du type d'échantillon (cellulaire ou tissulaire) analysé varie fortement en fonction des études. Ce choix dépend en partie de l'objectif de l'étude. En effet, si l'objectif est de mieux comprendre la physiopathologie de la maladie, il est préférable d'étudier les organes ou tissus cibles de la maladie. En revanche, si l'objectif est d'identifier des bio-marqueurs diagnostiques ou thérapeutiques, les échantillons faciles à obtenir tels que le sang total sont à privilégier.

Dans la SpA, l'enthèse constitue l'une des meilleures cibles puisque son inflammation fait partie des atteintes typiques de la pathologie. Toutefois, il s'agit d'un matériel difficile à obtenir et son profil d'expression génique n'a pas encore été étudié dans la SpA. Certaines études se sont néanmoins intéressées au matériel articulaire, notamment au liquide synovial provenant des articulations sacro-iliaques<sup>399</sup>, à des biopsies synoviales de genou<sup>394,397</sup> ou aux ligaments des articulations coxo-fémorales<sup>398</sup>. Laukens *et al.* ont quant à eux choisi d'étudier des biopsies coliques<sup>392</sup>. Toutefois, ce type d'échantillon est souvent difficile à obtenir et la plupart des études disposaient de faibles effectifs, ce qui limite leur puissance statistique.

Parmi les échantillons plus facilement accessibles, les cellules sanguines appartenant au système immunitaire constituent un choix intéressant étant donné leur implication possible dans des maladies inflammatoires. Plusieurs études ont comparé le transcriptome des cellules mononuclées du sang périphérique (peripheral blood mononuclear cells ou PBMCs) entre individus atteints de SpA et sujets sains<sup>387,388,390</sup>. Smith *et al.* ont quant à eux analysé les macrophages dont plusieurs études avaient préalablement souligné le rôle potentiel dans la pathogénie de la SpA<sup>396</sup>. Ces cellules n'étaient toutefois pas des cellules primaires, mais ont été obtenues *ex vivo* à partir de monocytes du sang circulant. Ce choix résultait de la difficulté à obtenir une quantité suffisante d'ARN pour une étude pangénomique à partir d'un compartiment cellulaire minoritaire du sang circulant. Il avait également l'avantage de soustraire les cellules de leur environnement et ainsi de limiter l'influence des traitements médicamenteux potentiellement en cours chez les patients.

Le sang total est à privilégier dans les études cherchant à identifier des biomarqueurs, afin de faciliter leur utilisation potentielle à grande échelle. Quatre études de ce type ont été publiées dans la SpA<sup>386,391,393,395</sup>. L'analyse des ARN obtenus à partir du sang total présente toutefois plusieurs difficultés techniques. En premier lieu, il est nécessaire de s'affranchir de l'interférence des ARN d'intérêt par les ARN messagers très abondants de la globine<sup>400</sup>. En second lieu, il faut garder à l'esprit que l'analyse du transcriptome à partir du sang total représente un mélange de cellules de différents types en pourcentage variable d'un individu à l'autre<sup>401</sup>. Cette hétérogénéité peut masquer des différences d'expression génique si elles concernent un sous-type cellulaire minoritaire.

### III.3.1.3 Choix du comparateur

Le choix des témoins auxquels sont comparés les sujets atteints de SpA constitue une différence supplémentaire entre les études. Le plus souvent, le groupe témoin était constitué d'individus sains ayant un sexe ratio et un âge moyen similaires à ceux des patients<sup>386–388,390,393,395,396</sup>. Néanmoins, les différences d'expression mises en évidence dans ce cas peuvent être en lien direct avec la pathogénie de la SpA ou refléter une inflammation systémique non spécifique.



Afin d'améliorer la spécificité des résultats, des témoins souffrant d'autres pathologies le plus souvent inflammatoires (comme la maladie de Crohn, la polyarthrite rhumatoïde, le lupus systémique ou la sclérodermie) ont été inclus dans certaines études<sup>386,388,390,392,394,397,398</sup>. Le choix de ces témoins a aussi été parfois dicté par la difficulté d'obtenir certains types d'échantillons chez des sujets sains pour des raisons éthiques (biopsie synoviale par exemple)<sup>394,397,398</sup>.

### III.3.2 Contribution des études du transcriptome à la compréhension de la spondyloarthrite

Les études du transcriptome dans la SpA ont abouti à l'identification de nombreux gènes différentiellement exprimés (**Tableau 13**). Très peu de ces gènes sont communs à plusieurs études. Ce faible recoupement s'explique par les différences importantes de conditions expérimentales entre les études, ainsi que par les effectifs souvent faibles favorisant le risque de faux-positifs. Il faut également souligner le caractère en général arbitraire du choix des gènes qui ont fait l'objet d'une validation ou d'une réplication. Ce choix a souvent été guidé par le rôle connu de ces gènes et leur lien potentiel avec la pathogénie de la SpA. Cela explique qu'on retrouve un nombre important de gènes en rapport avec l'inflammation et le système immunitaire.

Afin de prendre au mieux en compte l'ensemble des gènes différentiellement exprimés, la liste de ces gènes a parfois été confrontée à des bases de données permettant de relier chacun des gènes à des fonctions ou des voies biologiques. Ainsi, grâce à l'utilisation du logiciel Ingenuity Pathways Analysis, une sur-représentation statistiquement significative des gènes de la voie JAK/STAT et de la voie de signalisation des TLR ont été mises en évidence respectivement par Duan *et al.* et Assassi *et al.*<sup>386,387</sup>.

La voie dont la dérégulation semble la mieux établie est celle de l'interféron gamma. En effet, parmi les 141 gènes différentiellement exprimés entre macrophages de SA et de témoins dans l'étude de Smith *et al.*, 78 étaient en lien avec l'interféron gamma<sup>396</sup>. La majorité de ces 78 gènes était sous-exprimée chez les patients. L'expression de l'interféron gamma lui-même était significativement diminuée chez les patients. Les différences observées disparaissaient après traitement des macrophages avec de l'interféron gamma. Cette dérégulation de la voie de l'interféron gamma a aussi été mise en évidence dans les cellules dendritiques de rats transgéniques pour HLA-B27, l'un des modèles animaux de SpA<sup>402</sup>.

Aucun profil d'expression univoque n'est donc ressorti des études du transcriptome dans la SpA. La faible puissance de la plupart de ces études limitant grandement leur interprétation, la réalisation d'une méta-analyse de ces études, notamment de celles analysant le même type d'échantillon, pourrait permettre d'obtenir des résultats plus significatifs<sup>403</sup>.

**Tableau 13 – Résultats des études du transcriptome dans la SpA**

PBMC : peripheral blood mononuclear cells (cellules mononucléées du sang circulant) ; SFMC : synovial fluid mononuclear cells (cellules mononucléées du liquide synovial) ; SI : sacro-iliaque ; SA : spondylarthrite ankylosante ; USpA : spondylarthrite indifférenciée ; T : témoin ; PR : polyarthrite rhumatoïde ; RP : rhumatisme psoriasique ; OA : arthrose ; NR : non réalisé

Étude	Type d'échantillon	Effectif		Nombre de gènes DE	
		Malades	Témoins	Découverte	Réplication
<b>Gu et al.</b> <sup>388</sup>	PBMC	7	6T 7PR 7RP	SpA vs T : 4 SpA vs PR : 3 SpA vs RP : 3	4/7
<b>Gu et al.</b> <sup>389</sup>	SFMC SI	5	6T	23	NR
<b>Rihl et al.</b> <sup>394</sup>	Biopsie synoviale	3	3	86	NR
<b>Laukens et al.</b> <sup>392</sup>	Biopsie colique	15	10	464	NR
<b>Smith et al.</b> <sup>396</sup>	Macrophages	8	9	141	NR
<b>Gu et al.</b> <sup>390</sup>	PBMC	28 USpA 21 SA	20	USpA vs T : 38 SA vs T : 6	6
<b>Sharma et al.</b> <sup>395</sup>	Sang total	18	25	1090	102
<b>Duan et al.</b> <sup>387</sup>	PBMC	18	18	452	3/4
<b>Haroon et al.</b> <sup>391</sup>	Sang total	16	0	Avant/après Infliximab : 1428	
<b>Assassi et al.</b> <sup>386</sup>	Sang total	16	14	83	2/2
<b>Pimentel-Santos et al.</b> <sup>393</sup>	Sang total	18	18	221	13/47
<b>Xu et al.</b> <sup>398</sup>	Ligament hanche	18	6	519	NR
<b>Thomas et al.</b> <sup>397</sup>	Biopsie synoviale	8 SpA	4T /30A	SpA vs T+OA : 416	NR

## OBJECTIFS DE LA THÈSE

---

La SpA est donc une maladie au fort déterminisme génétique. De gros efforts ont été déployés ces dernières années pour identifier les facteurs génétiques associés à la maladie en dehors du HLA-B27. Ces efforts se sont jusqu'à présent concentrés sur le sous-type SA et ont abouti à l'identification de plus de 20 locus de susceptibilité, principalement grâce à des approches de type GWAS. Ces approches ont toutefois montré leurs limites, puisque l'ensemble des facteurs identifiés n'expliquent qu'une faible portion de l'héritabilité totale de la maladie et que son architecture génétique reste en grande partie inexpliquée.

L'identification de nouveaux facteurs de susceptibilité passe par l'utilisation d'autres approches. Mon équipe d'accueil a tout d'abord fait le choix de travailler sur les SpA dans leur ensemble afin de mieux prendre en compte la complexité du phénotype. De plus, nous avons fait l'hypothèse qu'une partie de l'héritabilité manquante était liée à des variants rares. Par conséquent, nous avons entrepris des études familiales, plus performantes que les GWAS dans l'identification de ce type de variant. Enfin, nous avons complété ces études de génétique dite « classique » par des études d'expression génique afin d'expliquer les conséquences potentielles des facteurs de susceptibilité mis en évidence.

Le travail présenté dans ce manuscrit s'est donc organisé autour de deux axes distincts, mais complémentaires, avec comme principal objectif une meilleure compréhension des facteurs génétiques impliqués dans la survenue de la SpA. D'une part, nous avons utilisé une approche génétique dite « classique » (**Articles 1, 2 et 3**). D'autre part, des études de génomique fonctionnelle ont été menées (**Article 4**). Le caractère complémentaire de ces deux approches est illustré par l'**Article 5** qui associe polymorphismes génétiques et expression génique.

Dans la première partie, nous avons étudié des familles multiplex de SpA. Le génotypage de plus de 260 000 SNPs chez plus de 900 individus provenant de familles multiplex a permis dans un premier temps de réaliser une analyse de liaison non-paramétrique pangénomique dont les résultats sont présentés dans l'**Article 1**. Nous avons ensuite effectué une étude d'association intra-familiale dont les résultats les plus intéressants ont fait l'objet d'analyses complémentaires dans deux populations familiales indépendantes, l'une française, l'autre nord-américaine (**Article 2**). Enfin, ces données familiales ont également été utilisées dans une étude gène-candidat testant l'association dans l'ensemble des sous-types de SpA d'un des polymorphismes du gène *IL23R* précédemment associé à la SA (**Article 3**).

La seconde partie de ce travail s'est intéressée à l'expression génique des cellules dendritiques dérivées de monocytes (MD-DCs) dans la SpA. Nous avons tout d'abord comparé le transcriptome de MD-DCs de patients à celui de témoins sains dans le but d'identifier des gènes différentiellement exprimés entre les deux groupes (**Article 4**). L'**Article 5** fait le lien entre les deux parties du travail avec l'étude de l'influence de certains polymorphismes d'*ERAP1* associés à la SpA sur l'expression du gène dans le même modèle de MD-DCs.

Deux autres études plus axées sur l'épidémiologie sont présentées dans les annexes. La première s'est intéressée à la prévalence de la SpA et du HLA-B27 dans la population française (**Annexe 1**). Dans la seconde, nous avons cherché à identifier les facteurs épidémiologiques, cliniques et génétiques associés à la présence d'une sacro-iliite radiologique dans la SpA (**Annexe 2**).

## **DEUXIÈME PARTIE : TRAVAUX ORIGINAUX**

# **Whole-genome Single Nucleotide Polymorphism-based Linkage Analysis in Multiplex Families of Spondyloarthritis Reveals a New Susceptibility Locus in 13q13.**

*Félicie Costantino, Emmanuel Chaplais, Tifenn Leturcq, Roula Said-Nahal, Ariane Leboime, Elena Zinovieva, Diana Zelenika, Ivo Gut, Céline Charon, Gilles Chiochia, Maxime Breban\*, Henri-Jean Garchon\**

*\*Les deux derniers auteurs ont contribué de manière équivalente à ce travail.*

Article en préparation.

## **Mise en évidence d'un nouveau locus de susceptibilité à la spondyloarthrite en 13q13 grâce à une étude de liaison génome-entier utilisant des polymorphismes mono-nucléotidiques dans des familles multiplex.**

La SpA est une maladie présentant une forte prédisposition génétique avec une hérédité complexe faisant intervenir de nombreux facteurs de susceptibilité. Outre l'antigène HLA-B27 qui est présent chez près de 80% des malades, plus d'une vingtaine de gènes de susceptibilité ont été identifiés ces dernières années, principalement grâce à des approches de type GWAS. L'ensemble de ces facteurs n'explique toutefois qu'une faible proportion de l'héritabilité totale à la maladie<sup>256</sup>. Dans le but d'identifier de nouveaux facteurs génétiques de susceptibilité à la SpA, nous avons entrepris une analyse de liaison génome-entier à partir du génotypage de 260 000 SNPs chez 914 individus appartenant à 143 familles multiplex. Parmi ces individus, 462 étaient atteints de SpA.

L'analyse non paramétrique nous a permis d'identifier une région significativement liée à la SpA en dehors du CMH en 13q13, ainsi que huit locus obtenant un score de liaison suggestif. L'analyse approfondie du locus 13q13 a permis d'estimer le pourcentage de familles liées à ce locus à

environ 30%, si on considère un modèle de transmission co-dominant avec des pénétrances de 30 et 60% en fonction du portage d'un ou deux allèles. Par ailleurs, la mise en évidence d'événements de recombinaison méiotique chez des patients appartenant aux familles liées a permis de restreindre la région d'intérêt à un intervalle d'environ 1,3 Mb contenant seulement 6 gènes. Aucune association significative n'a été détectée entre les polymorphismes de cette région et la SpA dans les familles les plus liées à ce locus.

Une liaison significative entre le locus 13q13 et la SpA est rapportée ici pour la première fois. Cependant, un score de liaison suggestif avait déjà été obtenu dans une précédente analyse de liaison réalisée par notre groupe (on ne peut toutefois pas considérer cette analyse antérieure comme indépendante de l'étude actuelle, puisque 91 familles sont communes aux deux études) et dans la méta-analyse des trois études de liaison déjà publiées dans la SpA<sup>260,262</sup>. Cette région a aussi été impliquée dans d'autres pathologies fréquemment associées à la SpA, comme la maladie de Crohn (où une liaison significative a été observée dans une étude mais non répliquée) ou le psoriasis (où des polymorphismes du gène *LHFP* appartenant à la région d'intérêt ont été associés à la maladie)<sup>404,405</sup>.

En outre, huit locus ont obtenu un score de liaison suggestif, parmi lesquels la région 9q33 correspondant au locus SPA2 précédemment identifié comme lié à la SpA. Considérant que le seuil de suggestivité correspond au risque statistique d'obtenir un seul faux-positif du fait du hasard sur l'ensemble du génome, il est logique d'envisager que parmi ces huit locus, certains soient réellement liés à la SpA.

L'identification d'un nouveau locus significativement lié à la SpA en 13q13 dans cette étude souligne donc l'intérêt de la réalisation d'études de liaison dans les maladies complexes. Une nouvelle étude dans un deuxième jeu de familles indépendant est actuellement en cours, pour nous permettre de mieux distinguer les éventuels locus d'intérêt parmi ceux ayant obtenu un score de liaison suggestif. Enfin, la région d'intérêt délimitée au sein du locus 13q13 fait actuellement l'objet d'un séquençage ciblé chez des individus appartenant aux familles les plus liés, afin d'identifier le(s) variant(s) génétique(s) à l'origine du signal de liaison.

**Whole-Genome Single Nucleotide Polymorphism-based Linkage Analysis in  
Multiplex Families of Spondyloarthritis Reveals a New Susceptibility Locus in 13q13.**

Félicie Costantino, M D,<sup>1,2,3</sup> Emmanuel Chaplais, M S,<sup>1,3</sup> Tifenn Leturcq, M D,<sup>1,3</sup>  
Roula Said-Nahal, M D,<sup>2</sup> Ariane Leboime, M D,<sup>2</sup> Elena Zinovieva, Ph D,<sup>1,3</sup>  
Diana Zelenika, Ph D,<sup>4</sup> Ivo Gut, Ph D,<sup>4</sup> Céline Charon, Ph D,<sup>4</sup> Gilles Chiocchia, Ph D,<sup>1,3</sup>  
Maxime Breban, M D, Ph D<sup>1,2,3\*</sup>, Henri-Jean Garchon, M D, Ph D,<sup>1,3,5\*</sup>

*1 – INSERM U987, UFR Simone Veil, Versailles-Saint Quentin University, France*

*2 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

*3 – Université Paris Diderot, Sorbonne Paris Cité, Laboratoire d'Excellence, Paris, France*

*4 – National Genotyping Center (CNG), Evry, France*

*5 – Genetics Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

\* The two last authors shared equal contribution

This work was supported by a grant from Agence Nationale de la Recherche (grant ANR 2010 GEMISA). Félicie Costantino and Tifenn Leturcq were supported by a grant from the Société Française de Rhumatologie (SFR).

Address reprint requests and correspondence to:

Maxime Breban, M.D., Ph.D., Rheumatology Division  
Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France  
Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865  
E-mail: [maxime.breban@apr.aphp.fr](mailto:maxime.breban@apr.aphp.fr)

Keywords: ankylosing spondylitis; spondyloarthritis; linkage analysis; genetics.

Word Count: 2,732



## ABSTRACT

**Objective:** Spondyloarthritis (SpA) is a chronic inflammatory disorder with high heritability but with complex genetics. Apart from HLA-B27, most of the underlying genetic factors remain to be identified. We conducted a whole-genome high density non-parametric linkage analysis to identify new genetic factors of susceptibility to SpA.

**Methods:** 914 subjects including 462 with SpA from 143 multiplex families were genotyped using Affymetrix 250K microarrays. After quality control, 189,368 single-nucleotide polymorphism (SNPs) were kept for further analyses. Both non-parametric and parametric linkage analyses were performed using Merlin software. Association was tested with Unphased.

**Results:** Non-parametric linkage analysis identified two regions significantly linked to SpA: the MHC ( $\text{LOD}_{\text{max}} = 24.77$ ) and a new 13q13 locus ( $\text{LOD}_{\text{max}} = 5.03$ ). Additionally, 8 loci achieved suggestive LOD scores, including the previously identified SPA2 locus at 9q33 ( $\text{LOD}_{\text{max}} = 3.51$ ). Parametric analysis supported a co-dominant model in 13q13 with a maximum heterogeneity LOD, “HLOD” score of 4.462 ( $\alpha = 0.305$ ). Identification of meiotic recombination events around the 13q13 linkage peak in affected subjects from the 43 best-linked families allowed us to map the disease interval between 38.753 and 40.040 Mb. Family-based association analysis of the SNPs inside this interval in the best-linked families identified a SNP near to *FREM2* (rs1945502) which reached a p-value close to statistical significance ( $p = 5.8 \times 10^{-4}$ ).

**Conclusions:** We report here for the first time a significant linkage between 13q13 and SpA. Identification of susceptibility factor inside this chromosomal region through targeted sequencing in linked families is underway.

## INTRODUCTION

Spondyloarthritis (SpA) is one of the most common forms of chronic inflammatory rheumatisms with an estimated prevalence of 0.43% in the Western adult population (1). It is characterized by axial and/or peripheral joint inflammation, often in association with extra-articular inflammatory features such as psoriasis, uveitis or inflammatory bowel disease (IBD). Depending on its clinico-radiological presentation, several subsets have been described: ankylosing spondylitis (AS), psoriatic arthritis, arthritis associated with IBD, reactive arthritis and undifferentiated SpA. Familial aggregation among these subsets has been established long ago suggesting shared genetic factors (2). This concept was reinforced by extensive analysis of a large panel of families with multiple cases of SpA, leading to the conclusion that all of the subtypes should be considered together in genetic studies (3–5).

Familial aggregation of SpA is high with a sibling recurrence risk ratio of 40 (6). Disease heritability, as estimated by twin studies in AS, exceeds 90% (7). The most important part of heritability comes from the HLA-B27 allele in the major histocompatibility complex (MHC), present in approximately 75% of the patients (1). However, there is strong epidemiologic evidence to suggest that other genes are involved. First, only 1 to 5% of HLA-B27 positive individuals develop SpA (1,8). Second, the disease risk is 5.6 to 16 times greater in HLA-B27 positive relatives of AS patients than in the general population (8). Third, in HLA-B27 twin pairs, the concordance rate was higher in monozygotic than in dizygotic ones (7,9,10). Recurrence-risk modeling based on familial data supports the hypothesis of an oligogenic model with between 3 and 9 genes operating in addition to HLA-B27 (11).

Several groups have identified genetic polymorphisms outside the MHC involved in AS susceptibility. The most recent findings resulted from genome-wide association studies (GWAS). Since 2007, three GWAS conducted in Caucasian population have led to the identification of 12 susceptibility loci outside the MHC (12–14). More recently, an international study performed in 10,619 AS cases and 15,145 controls, using the ImmunoChip array especially designed for immunogenetics studies, has uncovered 13 additional loci (15). However, these newly discovered loci explain only a small additional fraction of AS heritability. In total, less than 25% of the heritability of AS was considered as explained (including 20% contributed by HLA-B27 and 5% by other loci) (15).

Most of the associations with AS discovered through GWAS involve common variants. It has been suggested that a significant part of the “missing heritability” in complex

diseases might be due to effects of rare variants that are poorly detected in GWAS (16). Linkage analyses could be more powerful than GWAS for the detection of such variants (17). In SpA, three genome-wide linkage studies have been previously published, two in AS and one in SpA as a whole (18–21). Only two loci besides the MHC reached significance threshold: one on 16q and the other on 9q31-34 (19,21). Here we report on a new genome-wide linkage analysis using a high-density panel of single nucleotide polymorphisms (SNP) in an extended set of multiplex SpA families.

## SUBJECTS & METHODS

### *Multiplex SpA families*

Caucasian multiplex SpA families, *i.e.* including more than one SpA case per family, were recruited through the Groupe Français d'Etude Génétique des Spondylarthropathies (GFEGS), as previously described (3). This study was approved by the local ethical committees of Cochin Hospital (Paris, France) and Ambroise Paré Hospital (Boulogne-Billancourt, France). Written informed consent was obtained from each participant.

The diagnosis of SpA was originally ascertained according to the internationally validated classification criteria of Amor and/or European Spondyloarthritis Study Group (22,23), as previously described (3). All anteroposterior radiographs of the pelvis were examined blindly and independently by 2 qualified examiners (RSN, MB), using an established grading system (24). Even if they were not developed at the time of recruitment, the Assessment of SpondyloArthritis international Society classification criteria for axial and peripheral SpA were applied *a posteriori* (25,26).

The family set consisted of 143 families including 914 genotyped subjects of whom 462 had SpA. Detailed clinical characteristics of the dataset are provided in **Table 1**. Pedigrees' structures are summarized in **Table 2**.

### *Genotyping*

A SNP-based genome scan was conducted using the Affymetrix 250K SNPs array that covers the entire genome with around 262,000 SNPs. Genotyping was performed according to the manufacturer's instructions by the Centre National de Génotypage (Evry, France).

Briefly, total genomic DNA was digested with the NspI enzyme, ligated to the adaptor and amplified by polymerase chain reaction (PCR). After purification of the PCR products, amplicons were quantified, fragmented, labeled and hybridized to the 250K SNPs mapping array. Genotyping calling was determined by the BRLMM algorithm (27).

#### *Genotyping quality control*

Gender corresponding to each DNA sample was checked by analysis of X chromosome heterozygosity using PLINK software (28). The relatedness of individuals was verified using the GRR program (29).

Quality control was also performed for each SNP. First, information data provided by Affymetrix was updated using dbSNP 135. A total of 5,587 SNPs were removed from analysis for various reasons (5,188 had more than 2 alleles, 303 had been deleted from dbSNP and 96 were double-hits). Physical position on the genome and allele phasing was also updated if necessary. Pedstats was used to detect Mendelian errors (30). SNPs were then selected according to the following parameters: genotyping rate >95%, minor allelic frequency >1% and a significant deviation from Hardy-Weinberg proportions ( $P > 1 \times 10^{-3}$ ). Subtle genotyping inconsistencies were also detected and removed with Merlin (31). Finally, 189,368 SNPs were kept for further analysis.

#### *Linkage analysis*

Multipoint non-parametric linkage was tested for all autosomal chromosomes using Merlin software (31) to calculate the Kong and Cox LOD score (32). Two genome-wide thresholds of LOD score were used according to Lander & Kruglyak recommendations: 3.65 for significance, 2.2 for suggestiveness (33).

Linkage disequilibrium (LD) between SNPs could inflate LOD score in linkage studies (34). To address this issue, we performed linkage analysis handling LD as implemented in Merlin (with  $r^2$  threshold of 0.1) in the best-linked regions.

Following non-parametric analysis, we additionally performed parametric linkage analysis of the best-linked loci under three generic models, dominant and recessive (each with an arbitrary 50% penetrance of the disease allele) and co-dominant (with 30% and 60% penetrance in the presence of one or two doses of the disease allele).

### *Haplotype reconstruction*

Regional haplotypes were reconstructed with Merlin, using a subset of SNPs selected on a minor allelic frequency  $> 0.2$  and pairwise  $r^2 > 0.8$ . LD for reconstructed haplotypes was determined using the Haploview program V4.2 (35). Haplotypes were visualized on family tree using HaploPainter (36).

### *Association analyses*

Family-based association study was performed using Unphased version 3.1.6 (37). Bonferroni correction was applied to determine significance threshold.

## **RESULTS**

### *Non-parametric linkage analysis*

Results of the whole-genome non-parametric linkage analysis in the whole dataset are summarized in **Figure 1** and **Table 3**. Two regions achieved significant LOD scores: the MHC ( $\text{LOD}_{\text{max}} = 24.77$ ) and the 13q13 region with a linkage peak at 39.7 Mb from the p-telomere ( $\text{LOD}_{\text{max}} = 5.03$ ) which represents a new hit. Additionally, 8 loci achieved suggestive LOD scores, notably including the previously identified SPA2 locus at 9q33 ( $\text{LOD}_{\text{max}} = 3.51$  at 130 Mb).

### *Detailed analysis of the new locus in 13q13*

Non-parametric linkage analysis of the chromosome 13 revealed a significantly linked region around 40 Mb from the p-telomere as shown in **Figure 2**. Linkage was still significant after taking in account LD (**Figure 2**).

A parametric analysis was then conducted to test the likelihood of generic models of the disease locus and to assess the proportion  $\alpha$  of linked families under these models. The co-dominant model (with 30% and 60% penetrance in the presence of one or two disease alleles, respectively) was best supported, with a maximum heterogeneity LOD (HLOD) score of 4.462 ( $\alpha = 0.305$ ). The peak HLOD score was 4.112 (with  $\alpha = 0.279$ ) under the dominant model and 0.731 (with  $\alpha = 0.076$ ) under the recessive model (**Supplementary figure 1**).

To further circumscribe the disease locus, we reconstructed haplotypes around the linkage peak in the 43 best-linked families having a ponderated LOD score  $\geq 0.16$  (i.e. a

threshold consistent with the  $\alpha$  coefficient of 0.3). In 23 of these families, we identified one haplotype segregating with the disease, whereas in the remaining ones, all the affected members shared both of their haplotypes. Moreover analysis of the 5 meiotic recombination events occurring in patients (an example of such recombination event is shown in **Supplementary figure 2**) allowed us to map the disease interval between 38.753 Mb and 40.040 Mb, consistent with the results of non-parametric and parametric linkage analysis (**Supplementary figure 3**). The SNPs of this interval were tested for association in the 43 best-linked families. We observed no significant association after Bonferroni correction. However, the SNP rs1945502, an intergenic SNP located in the upstream region of *FREM2*, yielded a p-value close to significance (nominal  $p = 5.8 \times 10^{-4}$ , corrected  $p = 0.08$ ) with an odds ratio of 2.3 (95% confidence interval: 1.3 to 4.0) (**Figure 3**). All the haplotypes segregating with disease in the best-linked families, but one, harboured the risk allele of this SNP.

### *The MHC*

As expected, the MHC region on 6p21, carrying the HLA-B27 alleles, yielded the highest LOD scores ( $\text{LOD}_{\text{max}} = 24.77$ ). More than 92% of the patients included in the study were HLA-B27 positive (**Table 1**) and a vast majority of the families were linked to this locus. Four families had no HLA-B27 positive affected member (one consisted of 4 SpA patients, another one of 3 cases and the 2 remaining contained 2 cases each). Only one of those families had no MHC haplotype cosegregating with the disease (**Supplementary figure 4**). In the three other ones, all the affected members shared at least one MHC haplotype, still consistent with MHC linkage and suggesting allelic heterogeneity. Considering the 14 families in the lower decile of the ponderated LOD score distribution in the MHC, we observed in 7 of them two HLA-B27 haplotypes coming from distinct unrelated members. In the other ones, at least one affected subject was HLA-B27 negative but there was no obviously recurring HLA-B allele.

### *Genetic heterogeneity*

Except for the MHC, we observed high linkage heterogeneity with an estimated fraction of 30% of the families linked to 13q13 and of 16% linked to the previously identified SPA2 locus in 9q33. We compared the LOD scores obtained by each family at the two non-MHC best-linked loci (13q13 and 9q33) and observed no correlation (pairwise Spearman's test:  $p = 0.73$ ;  $\rho = 0.03$ ). (**Supplementary figure 5**).

## DISCUSSION

This high density genome-wide linkage analysis performed in a large set of multiplex families provides evidence for the existence of a new non-MHC susceptibility locus for SpA on 13q13. A suggestive linkage signal was also observed in this same region in the previous genome scan performed by our group, using microsatellite markers ( $NPL_{\max} = 2.55$ ;  $LOD_{\max} = 1.41$ ) (21).

The present study however cannot be considered as a replication because of the overlap between both family sets. Hence, 47 of the 65 families included in the initial microsatellite genome scan were also part of the present study. In the pooled analysis of the 3 previously published linkage analyses in SpA, conducted by Carter *et al.* (38), the microsatellite marker D13S218 in 13q13 achieved moderate evidence of linkage ( $P < 0.01$ ) with suggestive linkage ( $P < 0.05$ ) in the GFEGS and North American Spondylitis Consortium (NASC) datasets (20,21). Besides, the 13q13 locus has also been shown as significantly linked to Crohn's disease (CD), a chronic IBD which is frequently associated with SpA (39). In our dataset, however, comparison of the patients belonging to 13q13-linked families with the others showed no significant increase in CD frequency (**Table 1**).

Further analyses of the 13q13 allowed us to delimit a 1.3 Mb disease interval mapped between 38.7 and 40.0 Mb. This interval contains 6 genes (UFM1, FREM2, STOML3, PROSER1, NHLRC3 and LHFP), 4 pseudogenes and 4 non-coding RNA genes. Among them, LHFP, UFM1 and FREM2 are potential candidates. Indeed, polymorphisms of LHFP (lipoma HMGIC fusion partner) were previously associated with psoriasis, a major extra-articular component of the SpA spectrum (40). However, we observed no increased psoriasis frequency in patients of 13q13-linked families (**Table 1**). Some of the functions of UFM1 may also be related to SpA pathogenesis. Indeed components of the Ufm1 cascade have been demonstrated to be induced specifically under endoplasmic reticulum stress (41,42), a process that has been implicated in SpA pathogenesis (43). Finally, FREM2 is known to encode a large integral membrane protein of 3,169 aa with numerous repeats of a chondroitin sulfate proteoglycan domain. It is expressed at the basement membrane of several epithelia. Mutations of FREM2 were described in patients affected with type 2 Fraser syndrome, a rare malformative recessive disorder. Concerning the other genes located in the susceptibility interval, there is either little information on their function in public databases or their known functions could not be easily related to SpA pathogenesis.

Except for the MHC, there was a poor overlap between the findings of our scan and those of the studies previously published by the Oxford group and NASC (19,20). Among the 8 suggestive loci, only two in 9q and 17q overlapped with or were close to those identified by the pooled analysis of Carter *et al.* (38). The 17q region also overlaps with the NPEPPS-TBKBPI-TBX21 gene cluster recently shown as associated with AS (15). Indeed, the suggestiveness threshold corresponds to statistical evidence for a hit expected to occur less than once at random in a genome-wide scan. Thus some of the suggestive loci we detected may be true positives.

Several explanations could be proposed to explain discrepancies between studies. First, there is probably a high level of genetic heterogeneity in SpA, as reported for most complex diseases (44). Such genetic heterogeneity may be increased by differences in the geographic origin of patients (for instance, French versus British or North American). However, genetic heterogeneity may be seen in an ethnically homogenous population. For instance, we estimated that only 30% of the families studied herein were linked to 13q13 while a smaller set (16%) was linked to 9q33.

Genetic heterogeneity might also depend on variations in the studied phenotype. Here, we included all subtypes of SpA and not only AS patients (requiring advanced radiographic sacroiliitis) as it was the case in the Oxford and NASC studies (19,20). Our choice was based on segregation analyses indicating that major shared genetic factors were expected to account for all SpA subtypes coexisting in families (3,4). On the other hand, some of the loci previously identified in studies restricted to AS could be related to structural damage, i.e. more to disease severity than causality. Hence, several reports suggested that structural damage in SpA is partly genetically determined (45,46) and we recently identified that the SNP rs11209026 in the *IL23R* gene was specifically associated with AS but not with SpA lacking radiographic sacroiliitis (47).

Another difference between this study and those previously published in the early 2000's concerns the type of genetic marker used (a dense array of roughly 190,000 effective SNPs here, and a limited number of around 370 microsatellites from the Applied Biosystems Prism Linkage Mapping Set Version 2.0, in the previous ones). It is likely that the chance of detecting linkage was maximized in the present study by using a dense map of markers that extracted the maximum inheritance information content (48). Although they are bi-allelic and thus less informative, SNPs are present at a far greater density than microsatellites throughout



the genome. Several studies have demonstrated that a dense map of SNPs has greater power to detect linkage than low density microsatellite maps (49–52). In some instances, the same dataset was genotyped using both microsatellites and SNPs and some significantly linked loci were identified only in the SNP-based analysis (50–52). Moreover, a high SNP density allows a more precise localization of disease locus (51,52).

To conclude, we report here for the first time a significant linkage between 13q13 and SpA, highlighting the interest of high density SNP-based genome scan in large pedigrees dataset. Sequencing of the region of interest delimited between 38.7 and 40.0 Mb from the p-telomere in patients from the best linked families could allow us to identify variants associated with the disease. Additionally, 8 loci (including the previously identified SPA2 locus) achieved genome-wide suggestiveness threshold. Replication study in another familial cohort would be expected to discriminate true linkage from false positivity.

## REFERENCES

1. Costantino F, Talpin A, Said-Nahal R, Goldberg M, Henny J, Chiocchia G, et al. Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann Rheum Dis* 2013. Dec 18. doi:10.1136/annrheumdis-2013-204436. [Epub ahead of print].
2. Moll JM, Haslock I, Macrae IF, Wright V. Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behcet's syndrome. *Medicine (Baltimore)* 1974;53:343–364.
3. Said-Nahal R, Miceli-Richard C, Berthelot JM, Duché A, Dernis-Labous E, Blévec G Le, et al. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. Groupe Français d'Etude Génétique des Spondylarthropathies. *Arthritis Rheum* 2000;43:1356–1365.
4. Said-Nahal R, Miceli-Richard C, D'Agostino MA, Dernis-Labous E, Berthelot JM, Duché A, et al. Phenotypic diversity is not determined by independent genetic factors in familial spondylarthropathy. *Arthritis Rheum* 2001;45:478–484.
5. Porcher R, Said-Nahal R, D'Agostino M-A, Miceli-Richard C, Dougados M, Breban M. Two major spondylarthropathy phenotypes are distinguished by pattern analysis in multiplex families. *Arthritis Rheum* 2005;53:263–271.

6. Dernis E, Said-Nahal R, D'Agostino M-A, Aegerter P, Dougados M, Breban M. Recurrence of spondylarthropathy among first-degree relatives of patients: a systematic cross-sectional study. *Ann Rheum Dis* 2009;68:502–507.
7. Brown MA, Kennedy LG, MacGregor AJ, Darke C, Duncan E, Shatford JL, et al. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum* 1997;40:1823–1828.
8. Linden SM van der, Valkenburg HA, Jongh BM de, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27:241–249.
9. Järvinen P. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis Rheum* 1995;38:381–383.
10. Pedersen OB, Svendsen AJ, Ejstrup L, Skytthe A, Harris JR, Junker P. Ankylosing spondylitis in Danish and Norwegian twins: occurrence and the relative importance of genetic vs. environmental effectors in disease causation. *Scand J Rheumatol* 2008;37:120–126.
11. Brown MA, Laval SH, Brophy S, Calin A. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann Rheum Dis* 2000;59:883–886.
12. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–1337.
13. Reveille JD, Sims A-M, Danoy P, Evans DM, Leo P, Pointon JJ, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010;42:123–127.
14. Evans DM, Spencer CCA, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–767.
15. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, Pointon JP, Robinson PC, Karaderi T, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* 2013;45:730–738.
16. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;40:695–701.
17. Bailey-Wilson JE, Wilson AF. Linkage analysis in the next-generation sequencing era. *Hum Hered* 2011;72:228–236.

18. Brown MA, Pile KD, Kennedy LG, Campbell D, Andrew L, March R, et al. A genome-wide screen for susceptibility loci in ankylosing spondylitis. *Arthritis Rheum* 1998;41:588–595.
19. Laval SH, Timms A, Edwards S, Bradbury L, Brophy S, Milicic A, et al. Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. *Am J Hum Genet* 2001;68:918–926.
20. Zhang G, Luo J, Bruckel J, Weisman MA, Schumacher HR, Khan MA, et al. Genetic studies in familial ankylosing spondylitis susceptibility. *Arthritis Rheum* 2004;50:2246–2254.
21. Miceli-Richard C, Zouali H, Said-Nahal R, Lesage S, Merlin F, Toma C De, et al. Significant linkage to spondyloarthropathy on 9q31-34. *Hum Mol Genet* 2004;13:1641–1648.
22. Amor B, Dougados M, Mijiyawa M. [Criteria of the classification of spondylarthropathies]. *Rev Rhum Mal Osteoartic* 1990;57:85–89.
23. Dougados M, Linden S van der, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. *Arthritis Rheum* 1991;34:1218–1227.
24. Linden S van der, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361–368.
25. Rudwaleit M, Heijde D van der, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–783.
26. Rudwaleit M, Heijde D van der, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
27. Anon. BRLMM: an Improved Genotype Calling Method for the GeneChip Human Mapping 500K Array Set. Available at: [[http://www.affymetrix.com/support/technical/whitepapers/brlmm\\_whitepaper.pdf](http://www.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf)].
28. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575.
29. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001;17:742–743.

30. Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 2005;21:3445–3447.
31. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30:97–101.
32. Kong A, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997;61:1179–1188.
33. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;11:241–247.
34. Huang Q, Shete S, Amos CI. Ignoring linkage disequilibrium among tightly linked markers induces false-positive evidence of linkage for affected sib pair analysis. *Am J Hum Genet* 2004;75:1106–1112.
35. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265.
36. Thiele H, Nürnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics* 2005;21:1730–1732.
37. Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008;66:87–98.
38. Carter KW, Pluzhnikov A, Timms AE, Miceli-Richard C, Bourgain C, Wordsworth BP, et al. Combined analysis of three whole genome linkage scans for Ankylosing Spondylitis. *Rheumatology (Oxford)* 2007;46:763–771.
39. Shugart YY, Silverberg MS, Duerr RH, Taylor KD, Wang M-H, Zarfes K, et al. An SNP linkage scan identifies significant Crohn's disease loci on chromosomes 13q13.3 and, in Jewish families, on 1p35.2 and 3q29. *Genes Immun* 2008;9:161–167.
40. Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet* 2008;4:e1000041.
41. Azfer A, Niu J, Rogers LM, Adamski FM, Kolattukudy PE. Activation of endoplasmic reticulum stress response during the development of ischemic heart disease. *Am J Physiol Heart Circ Physiol* 2006;291:H1411–1420.
42. Zhang Y, Zhang M, Wu J, Lei G, Li H. Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS ONE* 2012;7:e48587.
43. Colbert RA, Tran TM, Layh-Schmitt G. HLA-B27 misfolding and ankylosing spondylitis. *Mol Immunol* 2014;57:44–51.

44. McClellan J, King M-C. Genetic heterogeneity in human disease. *Cell* 2010;141:210–217.
45. Brophy S, Hickey S, Menon A, Taylor G, Bradbury L, Hamersma J, et al. Concordance of disease severity among family members with ankylosing spondylitis? *J Rheumatol* 2004;31:1775–1778.
46. Ward MM, Hendrey MR, Malley JD, Learch TJ, Davis JC Jr, Reveille JD, et al. Clinical and immunogenetic prognostic factors for radiographic severity in ankylosing spondylitis. *Arthritis Rheum* 2009;61:859–866.
47. Kadi A, Costantino F, Izac B, Leboime A, Said-Nahal R, Garchon H-J, et al. Brief report: the IL23R nonsynonymous polymorphism rs11209026 is associated with radiographic sacroiliitis in spondyloarthritis. *Arthritis Rheum* 2013;65:2655–2660.
48. Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995;57:439–454.
49. Kruglyak L. The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* 1997;17:21–24.
50. Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, et al. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am J Hum Genet* 2004;74:886–897.
51. Vieland VJ, Walters KA, Azaro M, Brzustowicz LM, Lehner T. The value of regenotyping older linkage data sets with denser marker panels. *Hum Hered* 2014;78:9–16.
52. John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, et al. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am J Hum Genet* 2004;75:54–64.

## TABLES & FIGURES

**Table 1.** Clinical characteristics of SpA patients included in the genome-wide linkage study

Characteristic	All (n = 462)	Linked to 13q13* (n = 134)	Not linked to 13q13** (n = 328)
Age in years, mean $\pm$ SD	47.1 $\pm$ 14.5	45.2 $\pm$ 14.3	47.9 $\pm$ 14.6
Disease duration in years, mean $\pm$ SD	23.1 $\pm$ 13.4	26.5 $\pm$ 18.0	23.5 $\pm$ 16.0
Sex ratio, % of men	51.3	54.5	50.0
<i>HLA-B27</i> positivity, %	92.4	93.3	92.1
Axial manifestations			
- Inflammatory back pain, %	98.1	98.5	97.9
- Radiographic sacroiliitis***, %	58.8	59.7	58.4
Peripheral manifestations			
- Peripheral arthritis, %	44.2	42.5	44.8
- Peripheral enthesitis, %	69.3	70.1	68.9
Extra-articular manifestations			
- Uveitis, %	27.9	25.4	29.0
- Psoriasis, %	25.5	25.4	25.6
- Inflammatory bowel disease, %	6.7	6.7	6.7
o Crohn's disease	3.5	3.7	3.35
o Ulcerative colitis	3.2	3	3.35
Classification criteria fulfillment			
- Amor, %	99.6	99.3	99.7
- ESSG, %	94.8	94.0	95.1
- ASAS	99.3	99.2	99.4
o axial, %	94.3	94.7	94.2
o peripheral, %	5.0	4.5	5.2

\* refers to patients belonging to families reaching a ponderated LOD score  $\geq 0.16$  in 13q13 locus;

\*\* refers to patients belonging to families having a ponderated LOD score  $< 0.16$  in 13q13 locus;

\*\*\* refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral

**Table 2.** Characteristics of the SpA multiplex families included in the study

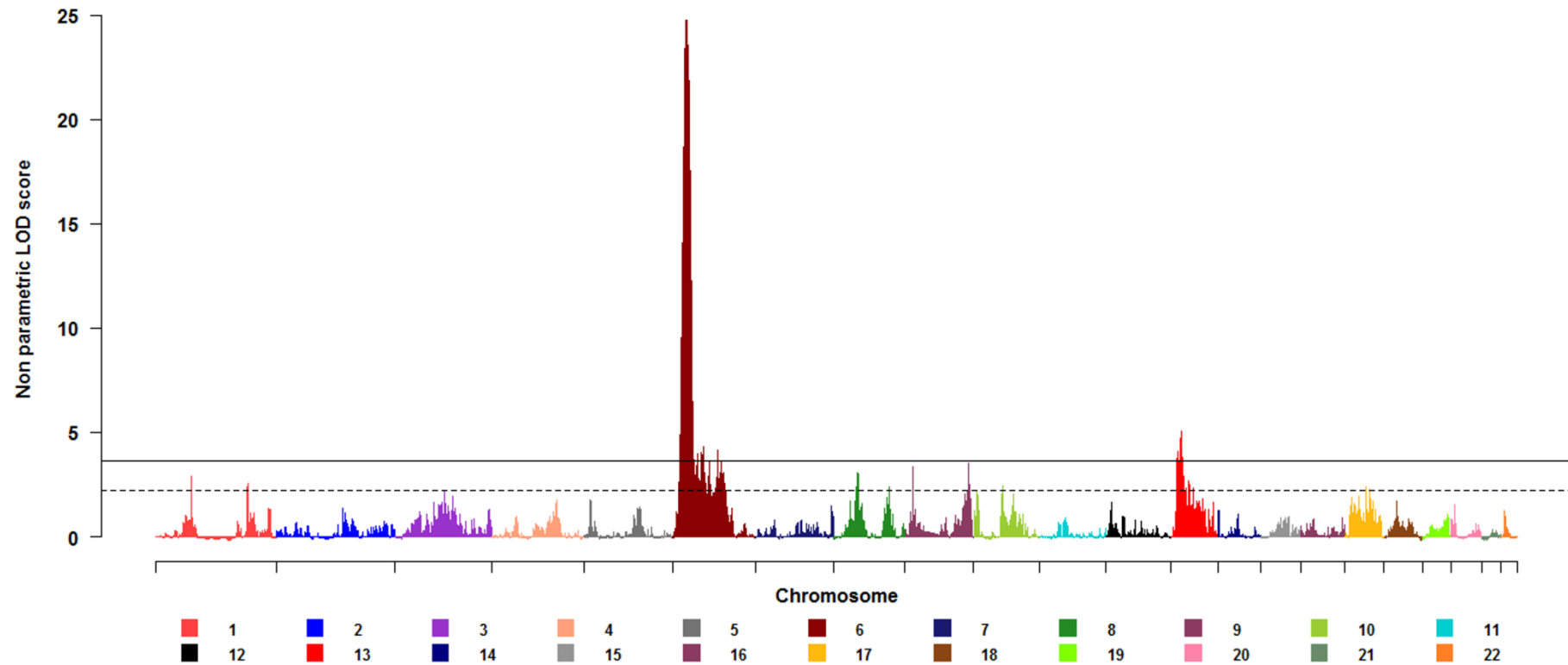
Characteristic	Family dataset (n =143)
Individuals per family	
- all, average no. (range)	7.57 (4 – 27)
- affected, average no. (range)	3.28 (2 – 11)
Generations per family, average no. (range)	2.47 (2 – 6)
Affected relative pairs no.	
- Sib-pairs	282
- Half-sibs	2
- Cousins	62
- Parent-child	163
- Grandparent-grandchild	10
- Avuncular	118

**Table 3.** Significant (bolded) and suggestive loci resulting from the non-parametric linkage analysis.

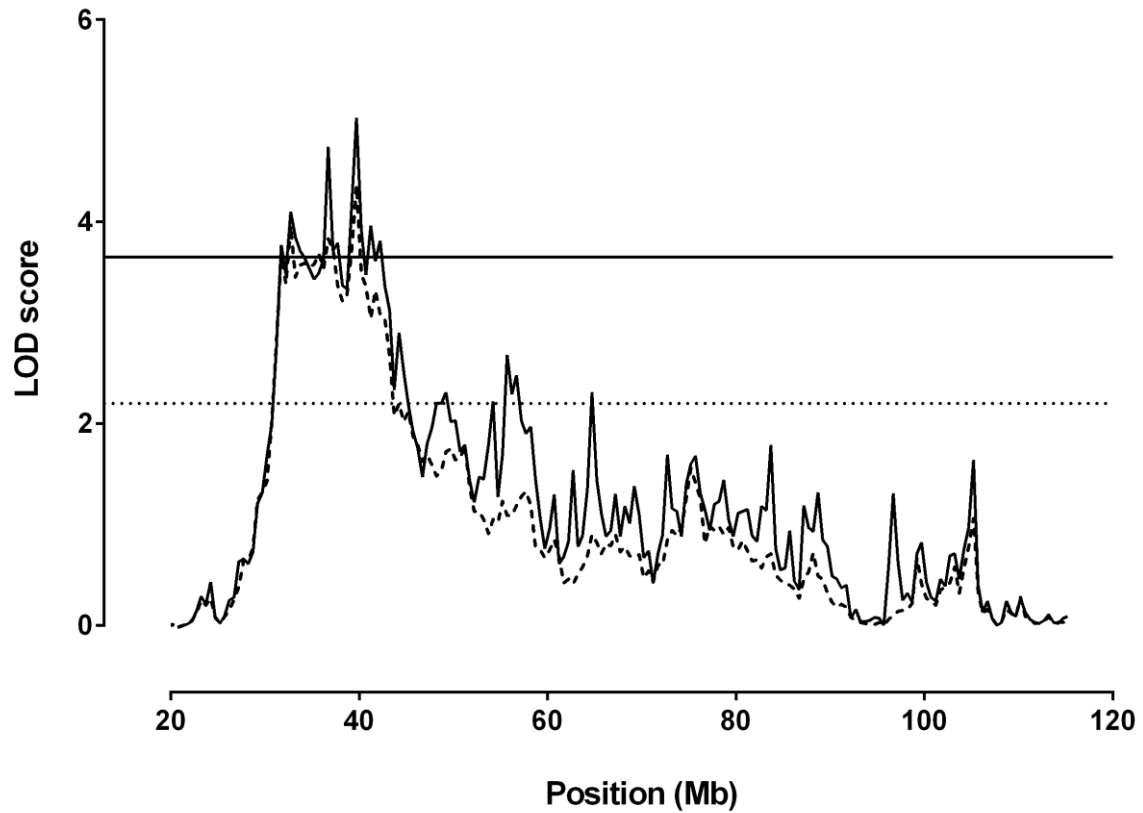
<b>Chromosome</b>	Position (Mb)	LOD score
1	73	2.87
1	189	2.55
<b>6</b>	<b>28.7</b>	<b>24.77</b>
8	48.2	3.06
8	114.2	2.4
9	15.7	3.34
9	130.2	3.51
10	58.7	2.44
<b>13</b>	<b>39.7</b>	<b>5.03</b>
17	44.0	2.37



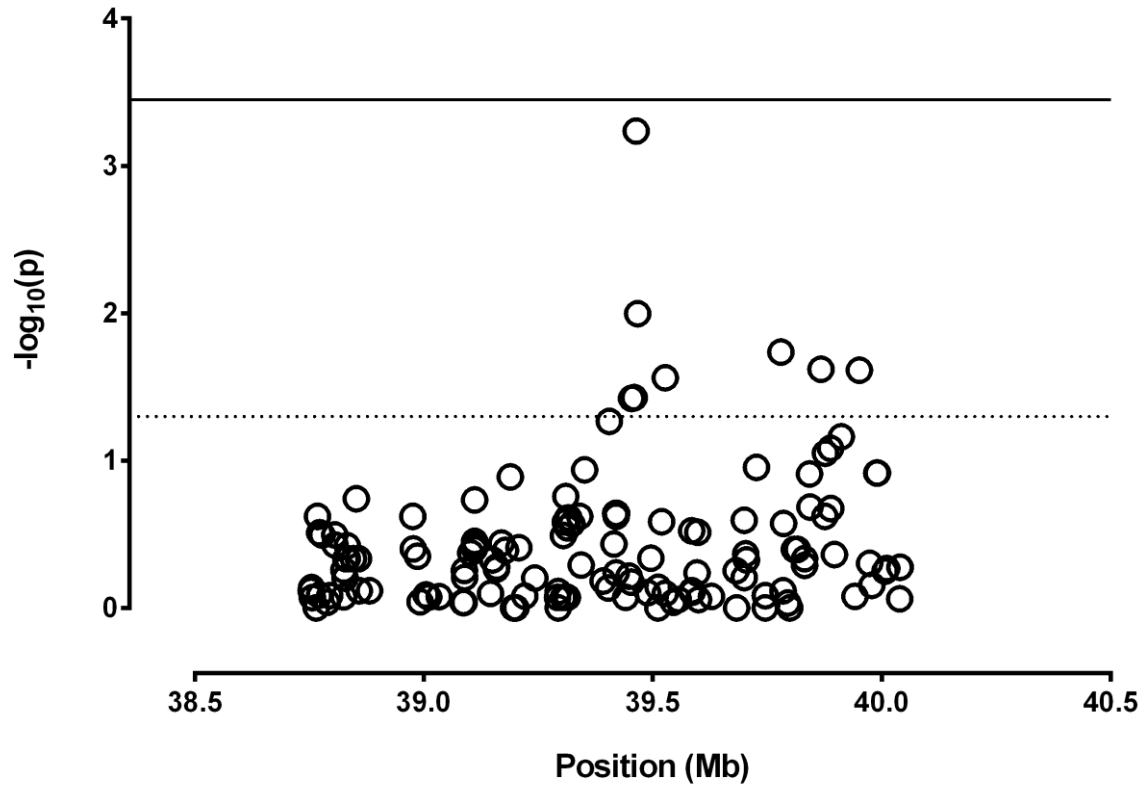
## FIGURES



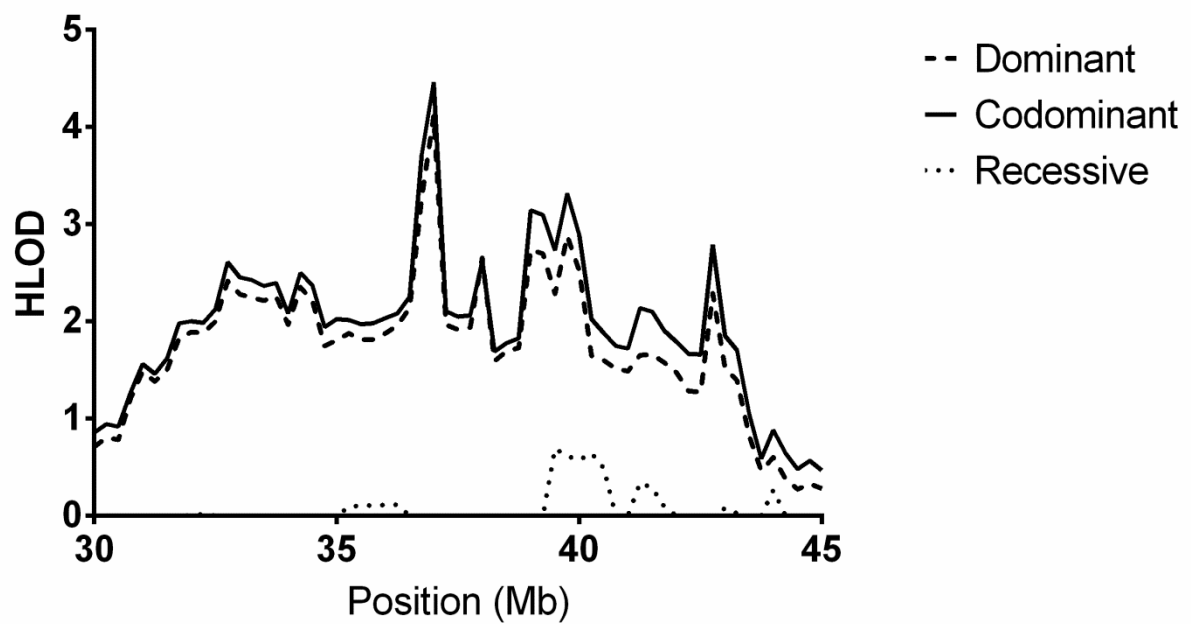
**Figure 1.** Genome-wide non-parametric linkage analysis results. The x axis indicates the marker positions on the genome with each chromosome represented with a distinct color. The y axis shows the Kong & Cox LOD score. Significant and suggestive LOD score thresholds are represented by solid and dashed line, respectively.



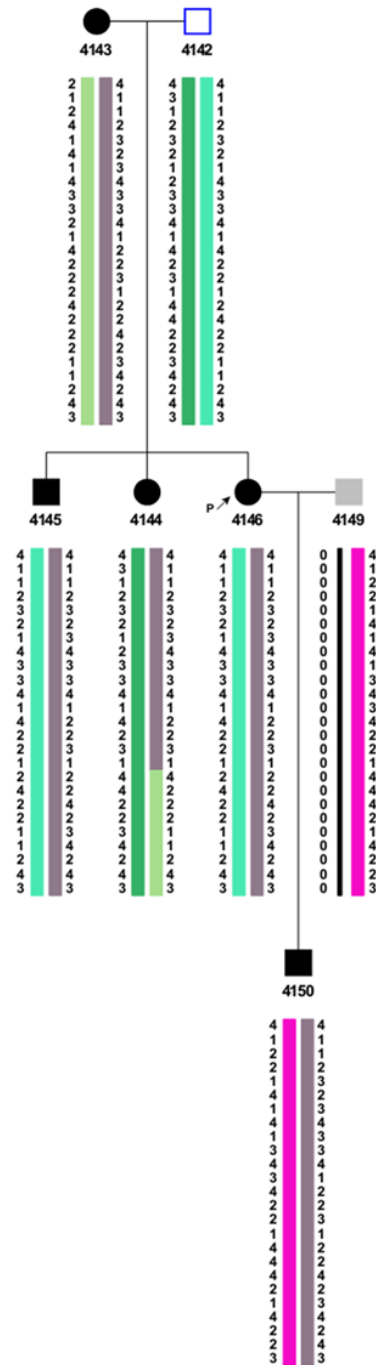
**Figure 2.** Non-parametric linkage analysis of chromosome 13, taking in account (dashed line) or not (solid line) the LD. The x axis represents the physical distance from the p-telomere and the y axis represents the Kong & Cox LOD score. Significant and suggestive LOD score thresholds are represented by solid and dotted horizontal lines, respectively.



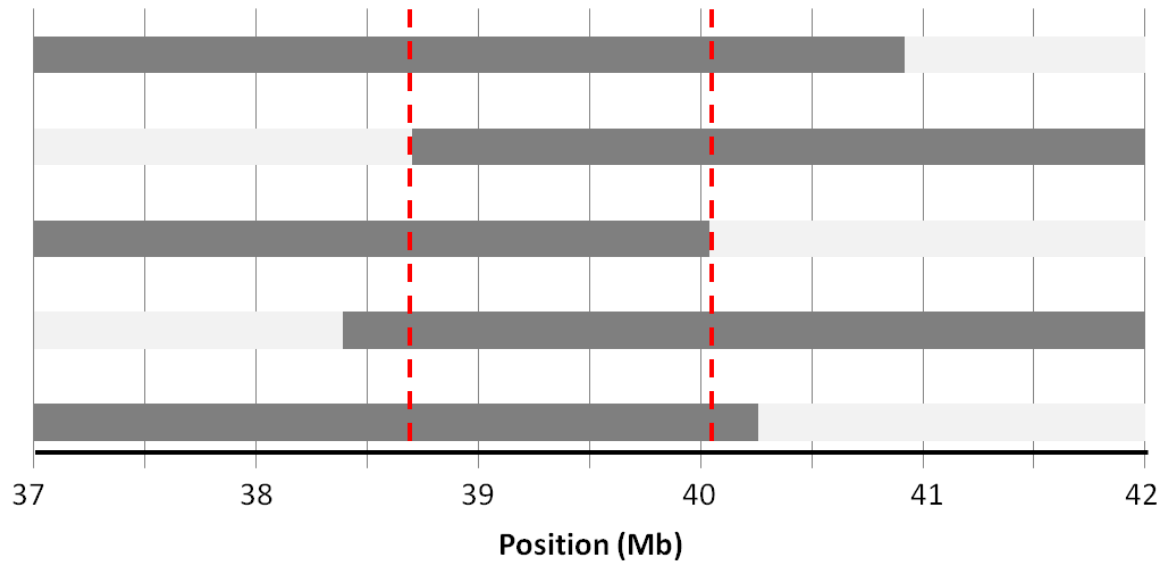
**Figure 3.** Family-based association study results. All the SNPs of the 13q13 disease interval have been tested for association in the 43 best-linked families (ponderated pLOD  $\geq 0.16$ ). The x axis represents position and the y axis represents the negative decimal logarithm of the p-value. Significant p-value threshold according to Bonferroni correction is represented by solid line and the dotted line corresponds to a nominal p-value of 0.05.



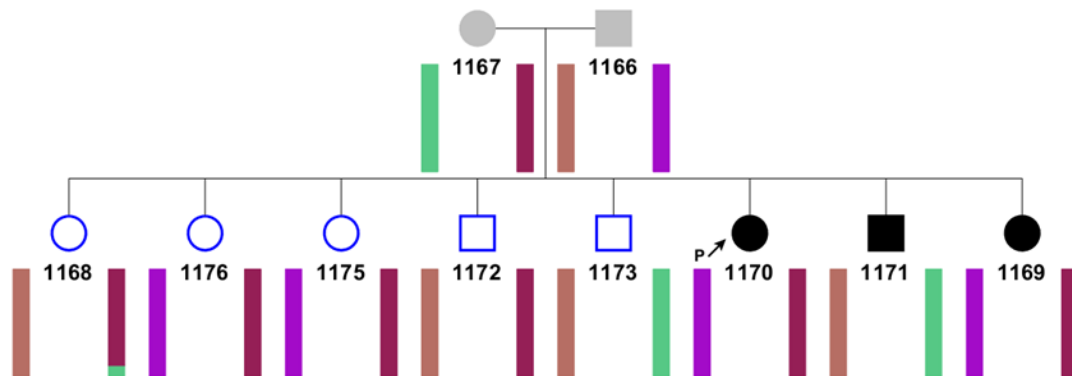
**Supplementary figure 1.** Parametric linkage analysis of chromosome 13 under a codominant (solid line), dominant (dashed line) or recessive (dotted line) model. The x axis represents the physical distance from the p-telomere and the y axis represents the HLOD score.



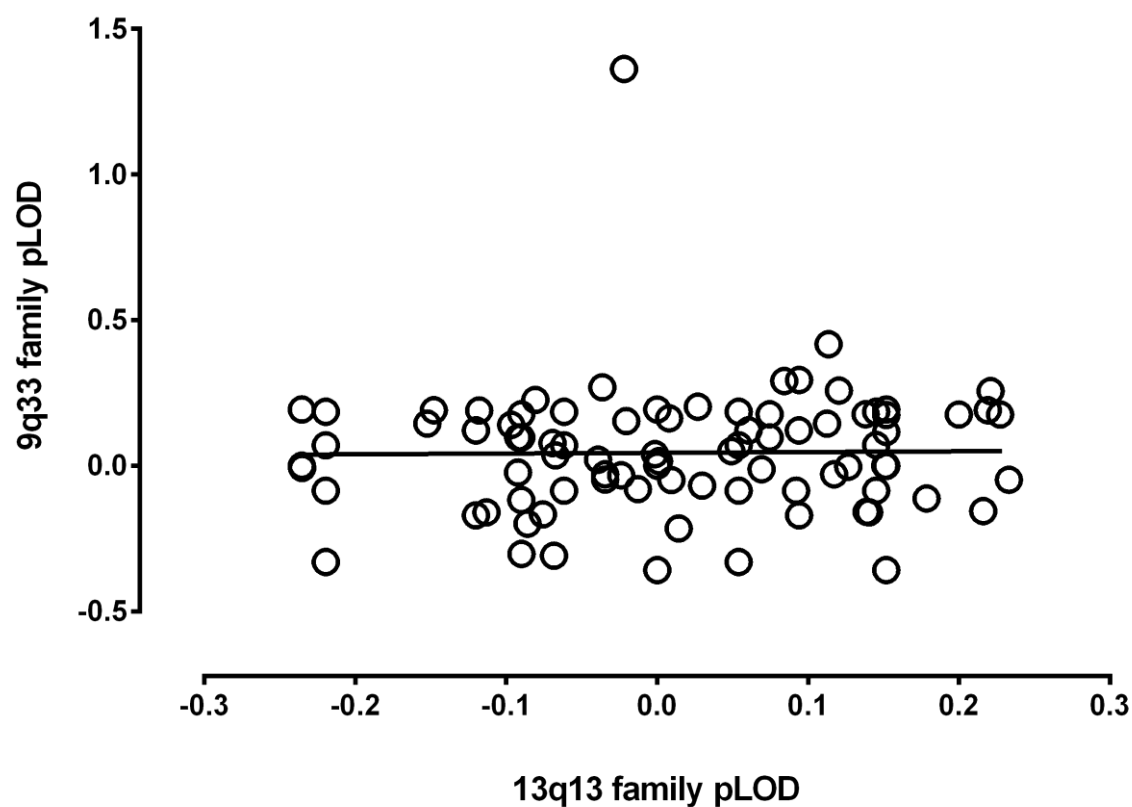
**Supplementary figure 2.** Example of an informative meiotic recombination event taking place at the 13q13 locus, in one linked pedigree. Haplotypes were reconstructed, using tag-SNP between 37 and 42 Mb from the chromosome 13 p-telomere. All the affected members of the pedigree harbor the grey-coloured haplotype with a recombination event in subject no. 4144. Closed symbols correspond to patients and open ones to nonaffected subjects. Shaded symbols indicate subjects who were not examined and genotyped. Their haplotypes were therefore deduced. SNP alleles are coded numerically.



**Supplementary figure 3.** Disease interval (framed in red) delimited by informative crossover<sup>2</sup> events in patients. For each informative crossover, the grey bar represents the portion of the recombinant haplotype which segregates with the disease in the other members of the family.



**Supplementary figure 4.** MHC haplotype reconstruction of the family no. 92. All the subjects (affected or not) are HLA-B27 negative. Subject no. 1171 does not share any haplotype with his two affected sisters (P: propositus).



**Supplementary figure 5.** No correlation between family pLODs at the new 13q13 locus and at the SPA2 locus in 9q33. For each family represented as a dot, the x axis represents the pLOD at 39.7 Mb from the p-telomere of chromosome 13 and the y axis represents the pLOD at 130.2 Mb from the p-telomere of chromosome 9. The black line represents the regression line.

### **A Family-Based Genome-Wide Association Study Reveals an Association of Spondyloarthritis with *MAPK14***

*Félicie Costantino, Emmanuel Chaplais, Alice Talpin, Brigitte Izac, Roula Said-Nahal, Ariane Leboime, Elena Zinovieva, Diana Zelenika, Ivo Gut, Céline Charon, John Reveille, Gilles Chiocchia, Henri-Jean Garchon\*, Maxime Breban\**

*\*Les deux derniers auteurs ont contribué de manière équivalente à ce travail.*

Article en préparation

#### **Mise en évidence d'une association de la spondyloarthrite avec *MAPK14* grâce à une étude d'association intra-familiale pangénomique**

Ce travail fait suite à celui présenté dans l'**Article 1** et poursuit le même objectif d'identifier de nouveaux facteurs de susceptibilité génétique à la SpA. Nous avons mené une analyse d'association intra-familiale pangénomique, en utilisant un jeu de familles quasi-identique à celui de l'étude précédente, génotypé à l'aide de puces de 260 000 SNPs.

Après réalisation du contrôle qualité, l'association de 230 801 SNPs avec la SpA a été testée au sein de ces familles multiplex. Si aucun SNP en dehors du CMH ne s'est significativement associé à la maladie en prenant en compte la correction de Bonferroni pour tests multiples, 42 obtenaient une valeur  $p < 10^{-4}$ . Pour tenter de confirmer ces premiers résultats, une étude d'extension portant sur les 357 meilleurs SNPs en dehors du CMH a été menée dans une cohorte indépendante de 121 familles françaises : 19 de ces SNPs étaient associés à la SpA au seuil nominal de  $p < 0.05$ . Une troisième cohorte de 192 familles nord-américaines a enfin été utilisée pour tester 20 des marqueurs les plus significatifs résultant des étapes précédentes, permettant ainsi la réplification d'une association significative pour trois de ces marqueurs. L'analyse combinée des trois cohortes portant



sur ces trois marqueurs a mis en évidence un score d'association proche de la significativité genome entier pour l'un des marqueurs, rs7761118.

Ce SNP est situé dans l'un des introns de *MAPK14* sur le bras court du chromosome 6 et qui code pour la p38 $\alpha$  MAP Kinase. Compte-tenu de sa localisation relativement proche du CMH, nous avons réalisé des analyses conditionnées sur l'allèle HLA-B27 permettant de nous affranchir du poids de ce facteur génétique. Une association indépendante du HLA-B27 ( $p < 0.05$ ) a été confirmée dans les trois cohortes. Ce gène est un bon candidat dans la survenue de la SpA, puisqu'il est impliqué dans des pathologies inflammatoires, donc certaines proches de la SpA comme la maladie de Crohn<sup>406</sup>.

Par ailleurs, nous nous sommes intéressés plus précisément aux SNPs situés dans des gènes connus pour être associés à la SA, d'après les résultats de précédents GWAS portant sur la SA. Une association nominale ( $p < 0.05$ ) était mise en évidence entre des polymorphismes à l'intérieur des gènes *IL23R*, *PTGER4*, *ERAP1*, *BACH2* and *ZMIZ1* et la SpA. Ces résultats renforcent le concept de terrain génétique partagé entre les différents sous-types de SpA.

Cette étude d'association intra-familiale suggère donc l'implication de *MAPK14* dans la pathogénie de la SpA. Toutefois, une étude plus précise de l'association des polymorphismes de ce gène avec la maladie devrait nous permettre d'identifier le(s) variant(s) causal(aux). Des études fonctionnelles seront par ailleurs nécessaires afin de mieux comprendre le rôle de *MAPK14* dans la SpA.

**Brief Report: A Family-Based Genome-Wide Association Study Reveals an Association  
of Spondyloarthritis with *MAPK14***

Félicie Costantino, M D,<sup>1,2,3</sup> Emmanuel Chaplais, M S,<sup>1,3</sup> Alice Talpin, Ph D,<sup>1,3</sup>  
Brigitte Izac, M S,<sup>1,3</sup> Roula Said-Nahal, M D,<sup>2</sup> Ariane Leboime, M D,<sup>2</sup>  
Elena Zinovieva, Ph D,<sup>1,3</sup> Diana Zelenika, Ph D,<sup>4</sup> Ivo Gut, Ph D,<sup>4</sup> Céline Charon, Ph D,<sup>4</sup>  
John Reveille, M D, Ph D,<sup>5</sup> Gilles Chiocchia, Ph D,<sup>1,3</sup> Henri-Jean Garchon, M D, Ph D,<sup>1,3,6</sup>  
Maxime Breban, M D, Ph D,<sup>1,2,3\* \*</sup>

*1 – INSERM U987, UFR Simone Veil, Versailles-Saint Quentin University, France*

*2 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

*3 – Université Paris Diderot, Sorbonne Paris Cité, Laboratoire d'Excellence, Paris, France*

*4 – National Genotyping Center (CNG), Evry, France*

*5 – Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center  
at Houston, Houston, Texas, USA*

*6 – Genetics Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

\* The two last authors shared equal contribution

This work was supported by grant from Agence Nationale de la Recherche (grant ANR 2010 GEMISA). Félicie Costantino was supported by a grant from the Société Française de Rhumatologie (SFR). Alice Talpin was supported by a grant from Arthritis Fondation Courtin.

Address reprint requests and correspondence to:

Maxime Breban, M.D., Ph.D., Rheumatology Division  
Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France  
Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865  
E-mail: maxime.breban@apr.aphp.fr

Keywords: spondyloarthritis, ankylosing spondylitis, GWAS, family-based association study, MAPK14

Word Count: 1,965

## ABSTRACT

**Objective:** Spondyloarthritis (SpA) is a chronic inflammatory disorder with high heritability but with complex genetics. More than 20 loci have already been associated to the disease but they explain only a minor fraction of total heritability. We conducted a family-based genomewide association study (GWAS) to identify new non-MHC genetic factors associated with SpA.

**Methods:** Nine hundred and six subjects from 156 French multiplex families, including 438 with SpA, were genotyped using Affymetrix 250K microarrays. After quality control, 230,801 single-nucleotide polymorphism (SNPs) were kept for further analyses. Association was tested with Unphased. The best-associated non-MHC SNPs were then genotyped in two independent familial cohorts, one from France and the other from North America (including 215 and 294 subjects, respectively) in order to replicate associations.

**Results:** Forty-three non-MHC SNPs yielded strong association signal with SpA in the discovery cohort ( $p < 10^{-4}$ ). In the extension studies, association was replicated at a nominal p-value of  $p < 0.05$  for 16 SNPs in the second cohort and for 3 SNPs in the third cohort. Pooled analysis identified an association close to genomewide significance between rs7761118, an intronic SNP of *MAPK14*, and SpA ( $p = 6.2 \times 10^{-7}$ ). Such association appeared as independent of HLA-B27.

**Conclusions:** We report here for the first time a family-based GWAS study on SpA and identified a highly associated polymorphism near *MAPK14*. Further analyses are needed to better understand genetic association between *MAPK14* and SpA.

## INTRODUCTION

Spondyloarthritis (SpA) is a frequent chronic inflammatory rheumatic disorder with a prevalence of 0.43% in the Western adult population (1). It is characterized by joint inflammation of the spine, frequently associated with extra-articular inflammatory features such as psoriasis, uveitis, or inflammatory bowel disease (IBD). Different entities are described according to the disease's clinical presentation. Although each of these entities is defined by specific characteristics they share several major clinical and imaging features, as well as genetic predisposition.

SpA heritability is high and involves several genetic factors, among which the HLA-B27 allele in the major histocompatibility complex (MHC). However, there is strong epidemiologic evidence to suggest that other genes are involved. For instance, only 1 to 5% of HLA-B27 positive individuals develop SpA whereas the disease risk is 5.6 to 16 times greater in HLA-B27 positive relatives of AS patients than in the general population (1,2). More than 20 susceptibility loci outside the MHC have already been identified, most often through genome-wide association studies (GWAS) (3-6). However, these newly discovered loci account for only a small fraction of the disease predisposition. In total, less than 25% of the heritability of SpA was considered as explained (including 20% contributed by HLA-B27 and 5% by other loci) (6).

In order to identify new genetic factors of susceptibility to SpA, we performed a family-based GWAS. Most compelling associations resulting from this first stage were then tentatively replicated in two independent cohorts of trios.

## PATIENTS AND METHODS

### *Ethics Statement*

This study was approved by the local ethical committees of Cochin Hospital (Paris, France), Ambroise Paré Hospital (Boulogne-Billancourt, France) and University of Texas Medical School (Houston, Texas, United States). Written informed consent was obtained from each participant.

## ***Study population***

Three family-based cohorts were used in this study. The discovery cohort consisted of 156 multiplex families of SpA (906 genotyped subjects of whom 438 had SpA) recruited throughout France by the Groupe Français d'Etude Génétique des Spondylarthrites (GFEGS), as previously described (7). The first replication cohort included 121 families (either simplex trios or multiplex families) belonging to the GFEGS collection with 423 genotyped subjects including 215 SpA patients. Finally, we used a second replication cohort of 658 subjects from 192 North American families, including 294 SpA cases.

Detailed clinical characteristics of the enrolled patients are provided in **Table 1**. The Assessment of SpondyloArthritis international Society (ASAS) classification criteria for SpA were fulfilled in 99.3% in the discovery cohort and in 98.1% and 100% in the first and second replication cohorts, respectively (8,9).

## ***DNA isolation and genotyping***

Genotyping of the first cohort was performed by the Centre National de Génotypage (Evry, France) using the Affymetrix 250K SNPs array (Affymetrix, Santa Clara, California, USA) that covers the entire genome with around 262,000 SNPs. The best-associated non-MHC SNPs in the first part of the study were further genotyped in the second cohort using an Illumina BeadXpress array (Illumina, San Diego, California, USA). 20 SNPs were finally genotyped in the third cohort using Taqman genotyping assay (Applied Biosystems, Foster City, California, USA)

## ***Statistical analysis***

### ***Genotyping quality control***

Quality control measures were *applied* to the genome-wide genotype data. Information data provided by Affymetrix was first updated using dnSNP 135. Then we removed Mendelian errors and subtle genotyping inconsistencies using Pedstats and Merlin softwares (10). Finally, we removed SNPs with low genotyping rate (<95%) or with significant deviation from Hardy-Weinberg proportions ( $P < 1 \times 10^{-3}$ ). A similar procedure was applied for the genotyping data from Illumina BeadXpress array and Taqman genotyping assays.

### *Family-based association analysis*

Family-based association analysis was carried out using likelihood-based association analysis implemented in Unphased 3.1.6, using relevant options for conditioning analyses on HLA-B27 allele where necessary (11). Each dataset was analyzed first independently and then pooled together. Bonferroni correction was applied to determine significance threshold.

## **RESULTS**

### *Family-based GWAS results*

After quality control, 230,801 SNPs were kept for analysis. Results of the whole genome analysis are summarized in **Figure 1**. Several SNPs in the MHC reached genomewide statistical significance, using a conservative Bonferroni correction (according to which the threshold was set to  $2.17 \times 10^{-7}$ ). The best non-MHC marker, rs4765413, located at 127.066 Mb from the p-telomere in an intergenic area of the chromosome 12 yielded a p-value of  $2.6 \times 10^{-6}$ . This level of significance exceeds the  $10^{-5}$ – $10^{-6}$  thresholds advocated for gene-based scans (12). Moreover, 42 SNPs outside the MHC gave a strong association signal with  $p < 10^{-4}$  (**Supplementary table 1**).

Besides, we observed nominal association ( $p < 0.05$ ) for several polymorphisms located near susceptibility factors previously identified through GWAS in *IL23R*, *PTGER4*, *ERAP1*, *BACH2* and *ZMIZ1* (**Supplementary table 2**).

### *Replication and pooled studies*

We then selected the 357 best-associated non-MHC SNPs from the discovery step. We tested them for association in a first replication cohort of French origin. After quality control, 327 SNPs could be analyzed for association. No SNP achieved significant threshold after Bonferroni correction ( $p < 1.5 \times 10^{-4}$ ). Nevertheless, sixteen SNPs obtained a nominal p-value of  $p < 0.05$  and 3 of them reached P-value  $< 10^{-5}$  in the combined analysis of the whole French cohort. The most associated SNP in the whole French cohort was again rs4765413, yielding a p-value of  $5.78 \times 10^{-7}$ , close to genomewide significance (**Supplementary table 3**).

To try to confirm in another cohort from North America the best associations resulting from the previous steps of this study, 20 SNPs were further genotyped, of which one was

removed after quality control. Three SNPs achieved a nominal p-value < 0.05: rs16883860, rs17057749 and rs7761118 (**Table 2**)

In the combined analysis of the 3 datasets, two intronic variants of *MAPK14* yielded a p-value close to genomewide significance: rs7761118 and rs16883860 ( $p = 6.2 \times 10^{-7}$  and  $p = 1.4 \times 10^{-6}$ , respectively; **Table 2**). As *MAPK14* and HLA-B are separated by only 5 Mb, we performed conditional analyses on HLA-B27 to offset the major effect of HLA-B27. In the 3 cohorts, the association with rs7761118 yielded significant nominal P-value when the test allele was on the non HLA-B27 chromosome ( $p = 0.007$ ,  $p = 0.03$  and  $p = 0.05$  respectively; combined  $p = 0.01$ ).

## DISCUSSION

In order to identify new variants associated with SpA, we used a family-based approach in multiplex families. Whereas several case-control GWAS have previously been reported in AS, to our best knowledge, this is the first family-based GWAS reported in SpA. Although they require more subjects to be genotyped to gain as much power, family-based designs have several advantages compared to population-based approaches (13). First, because of familial aggregation, the studied sample is likely enriched for genetic factors that contribute to the disease risk. Second, family-based analyses are robust to the effects of population substructures such as slight differences in ethnicity. Third, detection of genotyping errors are more accurate in families especially for Mendelian errors.

As expected, the most associated polymorphisms in the discovery cohort belonged to the MHC. Outside of MHC, none of the SNPs reached genomewide significance level after Bonferroni correction. However, we could extend the study of the best-associated non-MHC SNPs in this discovery cohort to two other independent family-based cohorts, one from France and the other from North America. The intergenic SNP rs4765413 on chromosome 12 almost reached genomewide significance in the pooled French cohort ( $p = 5.78 \times 10^{-7}$ ). However, it was not replicated in the North American one ( $p = 0.46$ ).

Most interestingly 2 SNPs in *MAPK14* yielded a combined p-value close to genomewide significance. Among them, rs7761118 was the best-associated variant with a combined p-value of  $p = 6.2 \times 10^{-7}$ . Such association signal appeared as independent of HLA-B27 located not very far away, albeit at a distance where linkage disequilibrium was not expected to extend, as attested by conditional analyses. This variant is mapped to an intron of *MAPK14* a gene that codes for the p38 mitogen-activated protein kinase, whose functions can

easily be related to SpA pathogenesis. Indeed, this kinase is activated by pro-inflammatory cytokines and has been involved in inflammatory bowel disease, a condition frequently associated with SpA (14). Moreover, *MAPK14* was shown to be deregulated in relation with HLA-B27 expression (15).

Besides discovering new findings, we were interested in replicating gene associations previously reported with AS through case-control GWAS. Interestingly, we observed nominal associations for polymorphisms in *IL23R*, *PTGER4*, *ERAP1*, *BACH2* and *ZMIZ1* reinforcing the evidence of shared genetic background between SpA as a whole, such as in the present study, and AS (16).

To conclude, we have identified, through a family-based GWAS, several polymorphisms potentially associated with SpA including two intronic variants in *MAPK14* on chromosome 6 whose association with AS was close to genomewide significance and appeared independent of HLA-B27 located 5.65 Mb upstream. Larger studies will be needed to expand this result further and to identify causal variant.

## REFERENCES

1. Costantino F, Talpin A, Said-Nahal R, Goldberg M, Henny J, Chiocchia G, et al. Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann Rheum Dis* 2013. Dec 18. doi:10.1136/annrheumdis-2013-204436. [Epub ahead of print].
2. Linden SM van der, Valkenburg HA, Jongh BM de, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27:241–249.
3. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–1337.
4. Reveille JD, Sims A-M, Danoy P, Evans DM, Leo P, Pointon JJ, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010;42:123–127.
5. Evans DM, Spencer CCA, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–767.



6. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, Pointon JP, Robinson PC, Karaderi T, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* 2013;45:730–738.
7. Said-Nahal R, Miceli-Richard C, Berthelot JM, Duché A, Dernis-Labous E, Blévec G Le, et al. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. Groupe Français d'Etude Génétique des Spondylarthropathies. *Arthritis Rheum* 2000;43:1356–1365.
8. Rudwaleit M, Heijde D van der, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–783.
9. Rudwaleit M, Heijde D van der, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
10. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30:97–101.
11. Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008;66:87–98.
12. Thomas D, Clayton D. Betting odds and genetic associations. *J Natl Cancer Ints* 2012;96:421–423.
13. Benyamin B, Visscher PM, McRae AF. Family-based genome-wide association studies. *Pharmacogenomics* 2009;10:181–190.
14. Waetzig GH, Seegert D, Rosenstiel P, Nikolaus S, Schreiber S. p38 mitogen-activated protein kinase is activated and linked to TNF-alpha signaling in inflammatory bowel disease. *J Immunol* 2002;168:5342–5351.
15. Sahlberg AS, Penttinen MA, Heiskanen KM, Colbert RA, Sistonen L, Granfors K. Evidence that the p38 MAP kinase pathway is dysregulated in HLA-B27-expressing human monocytic cells: correlation with HLA-B27 misfolding. *Arthritis Rheum* 2007;56:2652–2662.
16. Baeten D, Breban M, Lories R, Schett G, Sieper J. Are spondylarthritides related but distinct conditions or a single disease with a heterogeneous phenotype? *Arthritis Rheum* 2013;65:12–20.

## TABLES

**Table 1.** Clinical characteristics of the SpA patients included in genome-wide linkage study

Characteristic	Discovery study (n = 438)	1 <sup>st</sup> replication (n = 215)	2 <sup>nd</sup> replication (n = 294 )
Age in years, mean $\pm$ SD	47.5 $\pm$ 14.9	41.2 $\pm$ 15.3	NA
Disease duration in years, mean $\pm$ SD	23.4 $\pm$ 13.7	23.3 $\pm$ 9.9	NA
Sex ratio, % of men	51.9	61.9	59.4
<i>HLA-B27</i> positivity, %	93.8	90.2	90.5
Axial manifestations			
- Inflammatory back pain, %	97.7	93.9	NA
- Radiographic sacroiliitis*, %	60.8	58.4	NA
Peripheral manifestations			
- Peripheral arthritis, %	44.4	39.2	NA
- Peripheral enthesitis, %	69.6	67.5	NA
Extra-articular manifestations			
- Uveitis, %	29.1	25.9	NA
- Psoriasis, %	25.2	31.6	NA
- Inflammatory bowel disease, %	6.6	6.6	NA
Classification criteria fulfillment			
- Amor, %	99.3	98.1	NA
- ESSG, %	94.7	98.1	NA
- ASAS	99.3	98.6	100
o axial, %	95.7	93.5	NA
o peripheral, %	3.6	5.1	NA

\* refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral

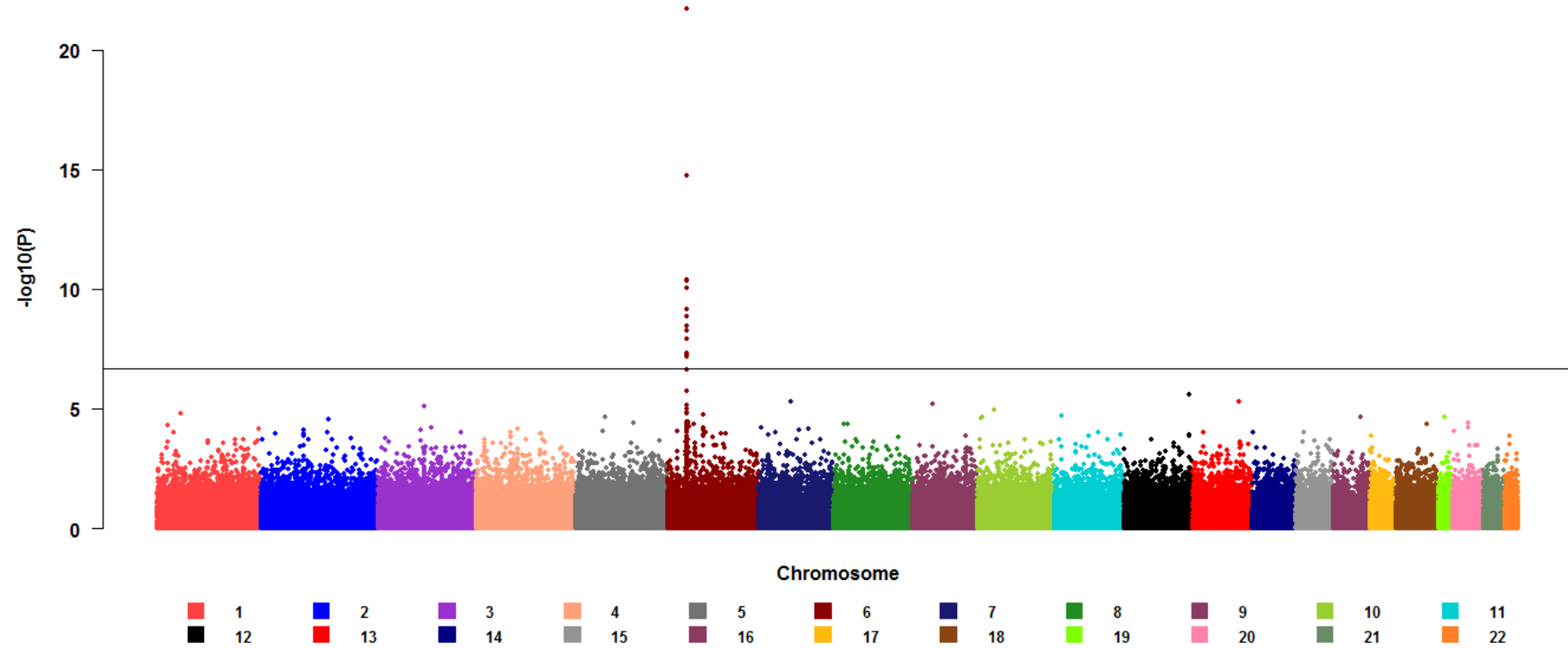
NA: not available.

**Table 2.** Most significant associations resulting from the combined family-based association analyses.

SNP name	Chr	Position (Mb)	Gene	Discovery cohort			First replication cohort			Second replication Cohort			Combined Cohort		
				T	U	P	T	U	P	T	U	P	T	U	P
<b>rs16883860</b>	6	36.002	<i>MAPK14</i>	0.08	0.12	0.0007	0.10	0.18	0.005	0.08	0.14	0.02	0.08	0.12	$1.4 \times 10^{-6}$
<b>rs7761118</b>	6	36.068	<i>MAPK14</i>	0.08	0.12	0.0007	0.06	0.012	0.009	0.08	0.15	0.006	0.08	0.12	$6.2 \times 10^{-7}$
<b>rs17057749</b>	13	60.739	<i>DIAPH3</i>	0.002	0.02	0.0009	0.03	0.07	0.009	0.02	0.04	0.03	0.01	0.03	$1.4 \times 10^{-5}$

Chr: chromosome; T, transmitted minor allelic frequency; U, untransmitted minor allelic frequency

## FIGURES



**Figure 1.** Manhattan plot of the genome-wide p-values of association. The x axis represents the position on the genome with each chromosome represented with a distinct color code. The y axis represents the negative decimal logarithm of p-values. Significant threshold assessed by Bonferroni correction is represented by solid line.

**Supplementary table 1.** Best associated on-MHC-SNPs in the discovery cohort

SNP name	Chromosome	Position (Mb)	Gene	T	U	P-value
rs7548252	1	34.794	-	0.28	0.36	5.23 x 10 <sup>-5</sup>
rs6671552	1	53.082	-	0.15	0.08	9.43 x 10 <sup>-5</sup>
rs10493390	1	66.313	<i>PDE4B</i>	0.04	0.15	1.58 x 10 <sup>-5</sup>
rs10927039	1	243.722	<i>AKT3</i>	0.08	0.16	6.64 x 10 <sup>-5</sup>
rs17517219	2	76.723	-	0.29	0.22	7.59 x 10 <sup>-5</sup>
rs16838470	2	138.012	<i>THSD7B</i>	0.01	0.07	9.42 x 10 <sup>-5</sup>
rs16843527	2	140.945	-	0.06	0.16	2.69 x 10 <sup>-5</sup>
rs17015351	3	77.614	<i>ROBO2</i>	0.05	0.14	8.24 x 10 <sup>-5</sup>
rs1003984	3	85.870	<i>CADM2-AS2</i>	0.21	0.15	7.88 x 10 <sup>-6</sup>
rs9813792	3	104.090	-	0.09	0.16	5.96 x 10 <sup>-5</sup>
rs9832487	3	165.921	-	0	0.05	9.42 x 10 <sup>-5</sup>
rs6857838	4	82.124	<i>PRKG2</i>	0.24	0.17	6.92 x 10 <sup>-5</sup>
rs41455	5	52.573	-	0.06	0.13	8.33 x 10 <sup>-5</sup>
rs2968222	5	56.928	-	0.03	0.07	2.18 x 10 <sup>-5</sup>
rs267303	5	114.587	<i>PGGT1B</i>	0.23	0.31	4.17 x 10 <sup>-5</sup>
rs9475771	6	13.915	-	0.42	0.34	8.50 x 10 <sup>-5</sup>
rs9462165	6	36.205	-	0.15	0.24	5.59 x 10 <sup>-5</sup>
rs9369631	6	46.708	-	0.15	0.25	4.60 x 10 <sup>-5</sup>
rs2262774	6	67.796	-	0.49	0.37	1.70 x 10 <sup>-5</sup>
rs2167984	6	67.819	-	0.46	0.35	5.98 x 10 <sup>-5</sup>
rs3734369	6	71.569	<i>SMAP1/B3GAT2</i>	0.02	0.05	9.42 x 10 <sup>-5</sup>
rs6976807	7	26.544	-	0.21	0.14	9.55 x 10 <sup>-5</sup>
rs17135796	7	53.573	-	0.01	0.07	5.11 x 10 <sup>-6</sup>
rs10224547	7	80.960	<i>AC004866</i>	0.15	0.09	7.75 x 10 <sup>-5</sup>
rs4729661	7	100.760	-	0.27	0.17	6.86 x 10 <sup>-5</sup>
rs1188311	8	13.070	-	0.03	0.15	4.21 x 10 <sup>-5</sup>
rs7005764	8	15.703	-	0.01	0.05	4.53 x 10 <sup>-5</sup>
rs1112066	9	23.244	-	0.03	0.15	6.52 x 10 <sup>-6</sup>
rs1901640	10	4.772	-	0.10	0.19	2.39 x 10 <sup>-5</sup>
rs11258455	10	6.435	-	0.02	0.06	2.18 x 10 <sup>-5</sup>
rs16923962	10	24.128	<i>KIAA1217</i>	0.02	0.10	1.17 x 10 <sup>-5</sup>
rs12575169	11	13.998	<i>SPON1</i>	0.14	0.22	1.98 x 10 <sup>-5</sup>
rs4765413	12	127.067	-	0.16	0.27	2.64 x 10 <sup>-6</sup>
rs1415245	13	34.749	-	0.10	0.07	9.75 x 10 <sup>-5</sup>
rs2321822	13	34.753	-	0.10	0.07	9.75 x 10 <sup>-5</sup>
rs427842	13	93.840	-	0	0.05	5.11 x 10 <sup>-6</sup>
rs16945685	16	77.461	<i>ADAMTS18</i>	0.01	0.06	2.18 x 10 <sup>-5</sup>
rs8095794	18	60.268	-	0.006	0.04	4.53 x 10 <sup>-5</sup>
rs1604432	19	31.262	-	0.06	0.10	2.15 x 10 <sup>-5</sup>
rs3088007	20	3.733	<i>HSPA12B</i>	0.27	0.37	8.69 x 10 <sup>-5</sup>
rs6515500	20	24.429	<i>GAPDHP53</i>	0.05	0.10	3.94 x 10 <sup>-5</sup>
rs6114732	20	24.431	<i>GAPDHP53</i>	0.05	0.10	6.37 x 10 <sup>-5</sup>

T, transmitted minor allelic frequency; U, untransmitted minor allelic frequency

**Supplementary table 2.** Associations results around genetic susceptibility factors associated with AS in the IGAS Immunochip study (6).

<b>Gene</b>	<b>Number of SNPs tested in the Affy250K array</b>	<b>Number of SNPs with p&lt;0.05</b>
<i>RUNX3</i>	5	0
<i>IL23R</i>	11	1
<i>GPR25 / KIF21B</i>	2	0
<i>PTGER4</i>	3	2
<i>ERAP1</i>	9	2
<i>IL12B</i>	4	0
<i>CARD9</i>	0	0
<i>LTBR / TNFRSF1A</i>	1	0
<i>NPEPPS / TBKBP1 / TBX21</i>	2	0
<i>IL6R</i>	1	0
<i>FCGR2A</i>	1	0
<i>UBE2E3</i>	13	0
<i>GPR35</i>	1	4
<i>BACH2</i>	26	5
<i>ZMIZ1</i>	21	0
<i>NAX2-3</i>	3	0
<i>SH2B3</i>	0	0
<i>GPR65</i>	4	0
<i>IL27 / SULT1A1</i>	0	0
<i>NOS2</i>	3	0
<i>TYK2</i>	0	0
<i>ICOSLG</i>	1	0

**Supplementary table 3.** Association results in the French family cohorts for the best-associated SNPs in the 1<sup>st</sup> replication step (nominal p < 0.05).

SNP name	Chr	Position (Mb)	Gene	Discovery cohort			Replication cohort			Combined cohort		
				T	U	P-value	T	U	P-value	T	U	P-value
<b>rs3768393</b>	1	226.881	<i>ITPKB</i>	0.02	0.05	0.0005	0.02	0.06	0.03	0.02	0.05	8.96 x 10 <sup>-5</sup>
<b>rs12105764</b>	2	40.646	<i>SLC8A1</i>	0.13	0.20	0.002	0.16	0.28	0.002	0.14	0.22	1.28 x 10 <sup>-5</sup>
<b>rs10865331</b>	2	62.551	-	0.48	0.43	0.07	0.52	0.53	0.03	0.50	0.43	0.005
<b>rs12374260</b>	4	47.537	<i>ATP10D</i>	0.41	0.52	0.0003	0.40	0.51	0.005	0.41	0.51	4.36 x 10 <sup>-6</sup>
<b>rs16883860</b>	6	36.002	<i>MAPK14</i>	0.08	0.12	0.0007	0.10	0.18	0.005	0.08	0.13	1.08 x 10 <sup>-5</sup>
<b>rs7761118</b>	6	36.068	<i>MAPK14</i>	0.08	0.12	0.0007	0.06	0.012	0.009	0.07	0.12	1.81 x 10 <sup>-5</sup>
<b>rs419574</b>	9	98.155	-	0.04	0.07	0.0006	0.004	0.3	0.004	0.03	0.06	2.74 x 10 <sup>-5</sup>
<b>rs1335405</b>	9	119.993	<i>ASTN2</i>	0.02	0.05	0.0005	0.04	0.09	0.008	0.02	0.06	2.37 x 10 <sup>-5</sup>
<b>rs10741628</b>	11	13.895	-	0.48	0.36	0.0004	0.48	0.39	0.03	0.48	0.37	4.77 x 10 <sup>-5</sup>
<b>rs4765413</b>	12	127.067	-	0.16	0.27	2.64 x 10 <sup>-6</sup>	0.22	0.32	0.02	0.19	0.28	5.78 x 10 <sup>-7</sup>
<b>rs17057749</b>	13	60.739	<i>DIAPH3</i>	0.002	0.02	0.0009	0.03	0.07	0.009	0.01	0.03	7.90 x 10 <sup>-5</sup>
<b>rs17082105</b>	13	67.685	<i>PCDH9</i>	0.006	0.03	0.0009	0.01	0.04	0.03	0.008	0.03	0.0002
<b>rs16951661</b>	13	96.844	<i>HS6ST3</i>	0.01	0.04	0.0004	0.003	0.03	0.02	0.007	0.03	4.60 x 10 <sup>-5</sup>
<b>rs11157326</b>	14	43.430	-	0.01	0.04	0.0004	0.09	0.15	0.02	0.04	0.06	0.0003
<b>rs12952507</b>	17	59.265	<i>ACCNI</i>	0.20	0.25	0.02	0.19	0.26	0.04	0.20	0.25	0.002
<b>rs6515500</b>	20	24.429	-	0.05	0.10	3.94 x 10 <sup>-5</sup>	0.05	0.09	0.04	0.05	0.10	7.31 x 10 <sup>-6</sup>

Chr: chromosome; T, transmitted minor allelic frequency; U, untransmitted minor allelic frequency

## Article 3

---

### **The *IL23R* Non Synonymous Polymorphism rs11209026 Is Associated With Radiographic Sacroiliitis in Spondyloarthritis**

*Amir Kadi\*, Félicie Costantino\*, Brigitte Izac, Ariane Leboime, Roula Said-Nahal, Henri-Jean Garchon, Gilles Chiochia, Maxime Breban*

*\*Les deux premiers auteurs ont contribué de manière équivalente à ce travail.*

Article publié dans « Arthritis & Rheumatism » en octobre 2013.

### **Le polymorphisme non synonyme de l'*IL23R* rs11209026 est associé à la sacro-iliite radiographique dans la spondyloarthrite.**

La première étude d'association génome-entier dans la SA, publiée en 2007, a mis en évidence une association significative entre la maladie et plusieurs polymorphismes du gène *IL23R*<sup>267</sup>. Cette association a depuis été répliquée dans plusieurs cohortes caucasiennes, notamment pour le polymorphisme non synonyme rs11209026 qui semble être le principal responsable du signal d'association. Ce SNP est également associé à la maladie de Crohn et au psoriasis, deux pathologies proches de la SpA<sup>405,407</sup>. L'objectif de cette étude était de déterminer si le SNP rs11209026 était associé à la SpA dans son ensemble.

Pour ce faire, le polymorphisme rs11209026 a été génotypé dans une cohorte française composée de 415 patients indépendants atteints de SpA et 312 témoins sains ainsi que dans une cohorte familiale de 383 trios (chacun constitué d'un individu malade et ses deux parents). Dans un premier temps, l'association a été testée entre ce polymorphisme et la SpA dans son ensemble dans chacune des cohortes. Une association significative n'a été observée que dans la cohorte familiale ( $p = 0,03$ , OR = 0,57). L'association a ensuite été testée en séparant les patients ayant une sacro-iliite radiologique selon les critères de New York modifiés de SA et ceux n'ayant pas de sacro-iliite radiologique (SpA non radiologique). Une forte association a été observée avec la SA dans les deux



cohortes ( $p = 4,5 \times 10^{-4}$  et  $p = 4,0 \times 10^{-3}$  respectivement). En revanche, aucune association n'a été détectée avec le groupe des patients sans sacro-iliite radiologique. On observait même une différence significative des fréquences alléliques entre les deux sous-groupes de patients (SA et SpA non radiologique ;  $p = 1,5 \times 10^{-3}$ ).

Une étude de corrélation phénotype/génotype a ensuite été réalisée grâce à une analyse de régression logistique. Cette étude a révélé que la présence d'une sacro-iliite radiologique ou un début précoce de la maladie, deux critères de sévérité de la maladie, étaient indépendamment associés à une fréquence plus faible de l'allèle protecteur A du SNP rs11209026. Comme précédemment rapporté, le taux sérique de CRP était associé à la sacro-iliite radiologique dans le sous-groupe de 209 patients pour lesquels une détermination de la valeur de CRP était disponible ( $p = 4.2 \times 10^{-4}$ ).

Ainsi, l'association du SNP rs11209026 avec la SpA est-elle restreinte aux formes avec sacro-iliite radiologique, au contraire de ce qui avait été précédemment observé pour les polymorphismes d'*ERAP1* comme rs30187<sup>383</sup>. Cela tend à montrer que certains facteurs génétiques influencent le niveau de susceptibilité à la maladie, alors que d'autres facteurs ont plutôt un rôle de modulateur phénotypique.

**Brief Report: The *IL23R* Non-Synonymous Polymorphism rs11209026 is Associated  
with Radiographic Sacroiliitis in Spondyloarthritis**

Amir Kadi, Ph D,<sup>1,2\*</sup> Félicie Costantino, M D,<sup>1,2\*</sup> Brigitte Izac, M S,<sup>1,2</sup>, Ariane Leboime, M D,<sup>3</sup>, Roula Said-Nahal, M D,<sup>3</sup>, Henri-Jean Garchon, M D, Ph D,<sup>1,2,4</sup> Gilles Chiocchia, Ph D,<sup>1,2</sup>  
Maxime Breban, M D, Ph D<sup>1,2,3</sup>

*1 – Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Paris, 75014, France;*

*2 – INSERM U1016 - Paris, 75014, France;*

*3 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), and Versailles Saint Quentin en Yvelines University, Boulogne-Billancourt, 92100, France;*

*4 – Genetics Division, Ambroise Paré Hospital (AP-HP), and Versailles Saint Quentin en Yvelines University, Boulogne-Billancourt, 92100, France;*

*\* The two first authors contributed equally to that work.*

This work was supported by grant from Agence Nationale de la Recherche (grant ANR 2010 GEMISA). Amir Kadi was supported by a grant from the Fondation pour la Recherche Médicale (FRM). Félicie Costantino was supported by a grant from the Société Française de Rhumatologie (SFR).

Address reprint requests and correspondence to:

Maxime Breban, M.D., Ph.D., Rheumatology Division  
Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France  
Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865  
E-mail: [maxime.breban@apr.aphp.fr](mailto:maxime.breban@apr.aphp.fr)

Keywords: ankylosing spondylitis; spondyloarthritis; single-nucleotide polymorphism; genetics; *IL23R*; sacroiliitis.

Word Count: 2,576

## ABSTRACT

**Objective:** Spondyloarthritis (SpA) is a group of inflammatory articular disorders sharing genetic background. The non-synonymous single-nucleotide polymorphism rs11209026 (*Arg381Gln*) in the *IL23R* gene, has reproducibly been shown as associated with ankylosing spondylitis (AS). Here, we examined the association between rs11209026 and SpA as a whole. Particular attention was devoted to genotype-phenotype correlation.

**Methods:** rs11209026 was genotyped in a French collection of 415 cases/372 controls and 383 trios. Association analysis was carried out in SpA as a whole group and then separately in AS and non-AS patients. Phenotype/genotype correlations were examined using a logistic regression analysis.

**Results:** A significant association between rs11209026 and SpA was only identified in the familial dataset ( $P = 0.03$ ; OR = 0.57). Strong association with AS was observed in both case/control and familial datasets ( $P = 4.5 \times 10^{-4}$  and  $4 \times 10^{-3}$ , respectively). In contrast, such association was not detected with the non-AS group. Furthermore, rs11209026 frequency was significantly different between the AS and non-AS patients ( $P = 2.5 \times 10^{-3}$ ). Phenotype/genotype correlation study revealed that both radiographic sacroiliitis and an early onset were independently associated with a lower frequency of the rare protective rs11209026 allele A in patients ( $P = 9 \times 10^{-3}$  and  $8 \times 10^{-3}$ , respectively).

**Conclusion:** Our study replicated the robust association between rs11209026 and AS, in the French population. However such association was restricted to AS patients, as compared to SpA without radiographic sacro-iliitis. The fact that it was independently conditioned by radiographic sacroiliitis and age at onset suggests that rs11209026 could affect the disease severity rather than susceptibility.

Spondyloarthritis (SpA) is the second most frequent form of chronic inflammatory rheumatic disorders with an estimated prevalence around 1% in adult Caucasian population (1). It comprises several closely related disorders characterized by axial and/or peripheral joint inflammation, often in association with extra-articular inflammation of the eye, skin or gut. Although each of these entities is defined by specific characteristics they share several major clinical and imaging features, as well as genetic predisposition (2).

SpA is a highly heritable disease involving multiple genetic factors, among which the HLA-B27 allele plays a prominent role (1). To discover other predisposing genetic factors, several genome-wide association studies (GWAS) have recently been conducted in AS, the prototypical form of SpA that relies on definite radiographic sacroiliitis for ascertainment. The first one revealed association with non-synonymous single-nucleotide polymorphisms (SNPs) in two non-MHC genes, *ERAP1* and *IL23R* (3). Since this original study, such associations have been consistently replicated in AS Caucasian populations (4). The most significantly AS-associated *IL23R* polymorphism across studies, rs11209026 (*Arg381Gln*) has also been implicated in susceptibility to inflammatory bowel diseases (IBD), psoriasis and psoriatic arthritis (5,6). In all these disorders, the minor allele A at rs11209026 (which is inferred to be non-ancestral) was strongly protective.

The *IL23R* gene is located on chromosome 1p31 and encodes one of both subunits of the receptor for interleukin-23 (IL-23), a cytokine critically involved in the expansion of Th17 cells. This subset of CD4<sup>+</sup> effector T lymphocytes, is characterized by the secretion of pro-inflammatory cytokines, such as IL-17A, IL-17F and IL-22 (7). Besides the association of *IL23R* polymorphisms with AS, several arguments indicate that both IL-23 and Th17 may play a role in the pathogenesis of SpA (8,9).

SpA is considered as a group of interrelated disorders sharing a common genetic background, as shown by their high familial aggregation and by their strong association with HLA-B27 (2). The validity of this unified concept was recently reinforced by the demonstration that polymorphisms in *ERAP1* and the IL-1 family gene cluster, already known as associated with AS were similarly associated with SpA lacking radiographic sacro-iliitis (10,11).

The aim of the present work was to determine whether rs11209026, a non-synonymous polymorphism in the *IL23R* gene also known to be associated with AS, would

similarly contribute to the whole SpA susceptibility and in particular in the absence of radiographic sacroiliitis

## **PATIENTS AND METHODS**

### ***Ethics Statement***

This study was approved by the local ethical committee of Cochin Hospital (Paris, France) and of Ambroise Paré Hospital (Boulogne-Billancourt, France). Written informed consent was obtained from each participant.

### ***Study population***

Genotyping was first performed in a case-control cohort consisting of 415 unrelated SpA cases and 372 healthy controls. For replication study, we used 383 trios (each consisting of 1 patient and both parents) belonging to 246 multiplex and 29 simplex families, which comprised overall 662 patients affected with SpA (the number of patients exceeded that of trios, because parents were frequently affected).

Caucasian families were recruited throughout France by the Groupe Français d'Etude Génétique des Spondylarthrites. In case/control panels, independent cases were recruited through the Rheumatology clinic of Ambroise Paré Hospital (Boulogne-Billancourt) or through the national self-help patients' organization: "Association Française des Spondylarthritiques". Independent controls were obtained from the "Centre d'Etude du Polymorphisme Humain", or were recruited as healthy spouses of cases.

Diagnosis was made according to the following internationally validated classification criteria: Amor's and/or European Spondylarthropathy Study Group criteria for SpA and modified New York criteria for AS (12). All anteroposterior radiographs of the pelvis were examined blindly and independently by 2 qualified examiners (RSN, MB), using an established grading system (12). Even if they were not developed at the time of recruitment, the Assessment of SpondyloArthritis international Society classification criteria for axial and peripheral SpA were applied *a posteriori* whenever possible (12,13). Detailed clinical characteristics of the enrolled patients are provided in Table 1.

### ***DNA isolation and genotyping***

Genomic DNA was extracted from peripheral blood using standard methods.

Genotypes of rs11209026 were obtained through the following techniques:

- Immunochip custom array (Illumina, San Diego, California, USA) for AS cases and controls, as part of the International Genetics of Ankylosing Spondylitis Consortium study (14);
- Roche LightCycler melting curve analysis using the the LightMix<sup>®</sup> Kit human *IL23R* Arg381Gln (TIB-MolBiol, Berlin, Germany; Cat.-No.40-0349-16) for non-AS cases;
- Affymetrix 250K SNP array (Affymetrix, Santa Clara, California, USA) for multiplex families;
- Illumina BeadXpress (Illumina, San Diego, California, USA) for simplex families;

Negative controls and duplicate samples were included to ensure accuracy of genotyping.

### ***Statistical analysis***

SNP was tested for deviation from Hardy–Weinberg equilibrium (HWE) separately in each control population. Disease association with single variants was assessed by Cochran-Armitage trend test for case-control study using PLINK 1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>). Family-based association analysis and pooled analysis were carried out using likelihood-based association analysis implemented in Unphased 3.1.6 (<http://unphased.sourceforge.net/>). A *P* value < 0.05 was considered as statistically significant. The Breslow-Day method was applied to calculate the homogeneity of odds ratio (OR) between both cohorts.

Three affection traits were considered successively: SpA as a whole, AS and SpA without radiographic sacroiliitis (referred hereafter as “non-AS” patients). Characteristics of both groups of cases are summarized in **Supplementary Table 1**.

This study had powers of 73% and 71% to identify an association between SpA and rs11209026, with a *P*-value of 0.05 and OR of  $\leq 0.6$ , in the case-control and family-based datasets, respectively.

To investigate the association of SNPs with patients clinical characteristics, we performed univariate logistic regression analysis. For each family, one SpA case was randomly selected and all the unrelated cases were also included in the analysis. To avoid bias in patients selection, we repeated the operation 10,000 times and reported *P*-value and OR distribution (median and interquartile range). Multivariate logistic regression models entered candidate variables (*P* value under 0.1 in univariate analysis).

Association of radiographic sacroillitis and of rs11209026 with elevated serum C-reactive protein (CRP) level (*i.e.* higher than laboratory positivity threshold) was tested by univariate logistic regression in the subgroup of patients for which serum CRP determination was available (*n* = 209).

## RESULTS

### *Case-control study*

Genotyping rate was 97.7%. Concordance between both genotyping methods used for this part of the study (ImmunoChip and LightCycler) was 99.3%. There was no significant deviation from HWE in the controls. Allele frequency of the minor allele A was similar in controls to those previously reported in Caucasian population and in HapMap samples of European descent.

There was no significant association between rs11209026 and the whole SpA group (*P* = 0.095). Association analysis was also conducted separately in patients with or without radiographic sacroillitis. Interestingly, rs11209026 was highly significantly associated with the AS subtype (OR = 0.33, 95% confidence interval (CI): 0.17 to 0.66, *P* = 4.5 × 10<sup>-4</sup>). In contrast, no association with this polymorphism was detected in the non-AS group (OR = 1.03, 95% CI: 0.63 to 1.68, *P* = 0.912) (**Table 2**). Furthermore, allelic frequencies were significantly different between the AS and non-AS patients (OR = 0.32, 95% CI: 0.15 to 0.68, *P* = 1.5 × 10<sup>-3</sup>) (**Table 2**).

### *Family-based study*

To replicate the foregoing results, we genotyped rs11209026 in 383 trios. Genotyping rate was 97.9%. There was no significant deviation from HWE in the control group.

The transmission disequilibrium test of rs11209026 in all SpA trios showed significant undertransmission of the minor allele of rs11209026 (OR = 0.57, 95% CI: 0.34 to 0.97,  $P = 0.028$ ) (**Table 2**). Again, separate analysis of AS and non-AS patients detected a strong association between the tested SNP and AS patients (OR = 0.27, 95% CI: 0.11 to 0.68,  $P = 4 \times 10^{-3}$ ) but no association in the non-AS group (OR = 1.13, 95% CI: 0.57 to 2.27,  $P = 0.723$ ) (**Table 2**).

### ***Pooled study***

We then combined analysis of case-control and family-based cohorts. Breslow-Day method identified no significant differences between the two datasets. Pooled analysis reinforced the significance of association between AS and rs11209026 (OR = 0.35, 95% CI: 0.20 to 0.59,  $P = 4.6 \times 10^{-6}$ ). No association was detected between this polymorphism and SpA as a whole or non-AS cases ( $P = 0.061$  and  $P = 0.41$ , respectively) (**Table 2**).

### ***Phenotype/genotype association analysis***

Phenotype/genotype correlation study revealed a significantly lower frequency of the rare protective rs11209026 allele A in patients with radiographic sacroiliitis ( $P = 9 \times 10^{-3}$ ) or with a younger age at onset ( $P = 4 \times 10^{-3}$ ). Each of these associations remained significant ( $P = 9 \times 10^{-3}$  and  $P = 4 \times 10^{-3}$ , respectively) in multivariate analysis (**Table 3**).

As previously reported, an elevated serum CRP level was significantly associated with sacroiliitis in the subgroup of 209 patients in which CRP had been determined (OR = 2.96, 95% CI: 1.62 to 5.42,  $P = 4.2 \times 10^{-4}$ ). Interestingly, carriers of the rs11209026 allele A had less frequently elevated CRP than carriers of the allele G (OR = 0.27, 95% CI: 0.09 to 0.86,  $P = 0.03$ ).

## **DISCUSSION**

Since initial GWAS discovery, numerous studies have confirmed that several *IL23R* polymorphisms were associated with AS. However, until now, none had investigated their association with SpA as a whole. We report here for the first time the results of an association study between rs11209026, the non-synonymous *IL23R* SNP most strongly associated with AS, and SpA patients with or without radiographic sacroiliitis. We could show that



rs11209026 was selectively associated with AS, either in sporadic or familial cases. We further identified an association between this SNP and both radiographic sacroiliitis, and age at disease onset.

Thus, association with rs11209026 seems to be restricted to AS. Indeed, in both of our datasets, we confirmed the protective effect of the rare allele A of this SNP on AS susceptibility, contrasting with its lack of association with SpA without sacroiliitis. Furthermore, the frequency of this SNP was significantly different between AS and non-AS patients. This result is in contrast with the confirmed association of *ERAPI* rs30187 with both AS and non-AS patients studied herein (data not shown), as previously reported (11). Thus, the lack of association observed with non-AS patients was unlikely to be explained by an insufficient power of the study. Frequency of the protective rs11209026 allele was lower in patients with radiographic sacroiliitis independently of their disease duration. Altogether these results suggest that in SpA, the rs11209026 allele A could exert a protective effect on the development of sacroiliac joint damage, including bony ankylosis.

IL-23, which signals through IL-23R complex, is one of the master regulators of Th17 effector T-cell subset. This subset is characterized by IL-17 and IL-22 production and is associated with strong pro-inflammatory responses. Several observations suggest a pivotal role for the IL-23/Th17 cells axis in the pathogenesis of SpA and related disorders, such as IBD and psoriasis. First, AS patients have elevated serum levels of IL-23 and IL-17 (9). Moreover, several genes involved in Th17 differentiation have been associated with these disorders (4). Finally, experimental animal models have implicated IL-23 and Th17 cells in SpA development (8,9).

The non-synonymous polymorphism rs11209026 is located in the initial portion of the IL-23R intra-cytoplasmic domain. By changing the highly conserved Arg381 for Gln381 at this position, the interaction between IL-23R and its signaling partner, Jak2 kinase, may be modified. The functional consequence of carrying the protective rs11209026 variant is a loss-of-function. Indeed, it was shown that CD4<sup>+</sup>CD45RO<sup>+</sup> and CD8<sup>+</sup> T cells from healthy Gln381 carriers have decreased IL-23-dependent IL-17 and IL-22 production and a lower percentage of circulating Th17 cells. Moreover cells transfected with the *IL23R* Gln381 variant exhibited decreased IL23-mediated signaling, as compared with those transfected with the *IL23R* Arg381 (15).

Given that the serum level of CRP is a strong positive predictor of radiographic sacroiliitis progression (16), we hypothesized that rs11209026 Gln381 might prevent radiographic sacroiliitis development by limiting the level of inflammation in SpA. Consistent with such hypothesis, an elevated level of CRP was associated with higher frequency of sacroiliitis in the present study. Moreover, we identified an association between rs11209026 and elevated CRP.

In our study, neither IBD nor psoriasis was associated with rs1109026 contrary to what is now well established (5,6). It may be explained by the insufficient power of this study especially relative to IBD (only 81 patients had such feature).

To conclude, the present study confirmed the strong association between the *IL23R* non-synonymous polymorphism rs11209026 and AS in the French population. We also demonstrated that such association could not be generalized to other SpA subtypes. The fact that it was independently conditioned by radiographic sacroiliitis and age at onset suggests that rs11209026 could affect disease severity rather than susceptibility, in particular by preventing structural damage.

## REFERENCES

1. Baeten D, Breban M, Lories R, Schett G, Sieper J. Are spondylarthritides related but distinct conditions or a single disease with a heterogeneous phenotype? *Arthritis Rheum* 2013;65:12–20.
2. Breban M, Miceli-Richard C, Zinovieva E, Monnet D, Said-Nahal R. The genetics of spondyloarthropathies. *Joint Bone Spine* 2006;73:355–362.
3. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–1337.
4. Reveille JD. Genetics of spondyloarthritis--beyond the MHC. *Nat Rev Rheumatol* 2012;8:296–304.
5. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science* 2006;314:1461–1463.

6. Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet* 2008;4:e1000041.
7. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang Y-H, et al. A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat Immunol* 2005;6:1133–1141.
8. Glatigny S, Fert I, Blaton MA, Lories RJ, Araujo LM, Chiochia G, et al. Proinflammatory Th17 cells are expanded and induced by dendritic cells in spondylarthritis-prone HLA-B27-transgenic rats. *Arthritis Rheum* 2012;64:110–120.
9. Sherlock JP, Joyce-Shaikh B, Turner SP, Chao C-C, Sathe M, Grein J, et al. IL-23 induces spondyloarthropathy by acting on ROR- $\gamma$ t+ CD3+CD4-CD8- enthesal resident T cells. *Nat Med* 2012;18:1069–1076.
10. Monnet D, Kadi A, Izac B, Lebrun N, Letourneur F, Zinovieva E, et al. Association between the IL-1 family gene cluster and spondyloarthritis. *Ann Rheum Dis* 2012;71:885–890.
11. Kadi A, Izac B, Said-Nahal R, Leboime A, Praet L Van, Vlam K de, et al. Investigating the genetic association between ERAP1 and spondyloarthritis. *Ann Rheum Dis* 2013;72:608–613.
12. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68 Suppl 2:ii1–44.
13. Rudwaleit M, Heijde D van der, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
14. Cortes A, Hadler J, Pointon J, Robinson PC, Karaderi T, Leo P, et al. Multiple novel loci harboring common and rare variants implicated in ankylosing spondylitis. *Nature genetics* 2013; in press.
15. Sarin R, Wu X, Abraham C. Inflammatory disease protective R381Q IL23 receptor polymorphism results in decreased primary CD4+ and CD8+ human T-cell functional responses. *Proc Natl Acad Sci USA* 2011;108:9560–9565.
16. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Märker-Hermann E, Zeidler H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. *Ann Rheum Dis* 2011;70:1369–1374.

## TABLES

**Table 1.** Clinical characteristics of spondyloarthritis patients in the study

Characteristic	Singleton cases (n=415)	Familial cases (n=662)
Age in years, mean $\pm$ SD	46.7 $\pm$ 13.1	46.1 $\pm$ 15.4
Age at onset in years, mean $\pm$ SD	26.2 $\pm$ 10.2	23.7 $\pm$ 9.1
Sex ratio, men:women (%)	208:207 (50.1/49.9)	365:297 (60/40)
<i>HLA-B27</i> positivity, %	78.8	93.5
Axial manifestations		
- Inflammatory back pain, %	97.3	97.6
- Radiographic sacroiliitis, % *	52.3	58.1
Peripheral manifestations		
- Peripheral arthritis, %	39.5	44.2
- Peripheral enthesitis, %	79.5	69.6
Extra-articular manifestations		
- Uveitis, %	24.3	28.5
- Psoriasis, %	27.6	28.3
- Inflammatory bowel disease, %	8.8	6.8
Spondyloarthritis subtype,		
- AS, %	52.3	58.3
- USpA, %	31.8	28.7
- PsA, %	13.3	10.9
- AIBD, %	2.6	1.7
- ReA, %	0	0.5
Classification criteria fulfillment		
- Modified New York, %	52.3	58.3
- Amor, %	97.7	98.9
- ESSG, %	99	96
- ASAS Axial SpA, %	83.4	95.1
- ASAS Peripheral SpA, %	11.3	4.3

Inflammatory bowel disease: Crohn's disease or ulcerative colitis; AS, ankylosing spondylitis; USpA, undifferentiated spondyloarthritis; PsA, psoriatic arthritis; AIBS, Inflammatory bowel disease-associated arthritis; ReA, Reactive Arthritis; ESSG, European Spondyloarthropathy Study group; ASAS, Assessment of SpondyloArthritis international Society;

The registered manifestations correspond to those present at the time of examination, or retrieved from past-medical history.

\* Refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral

**Table 2.** Results of the case/control and family-based allelic association studies for rs11209026

	Case-Control study					Family-based study					Pooled study	
	N	MAF	MAF	OR	P	N	T	U	OR	P	OR	P
		Patients	Controls	(95% CI)		Trios			(95% CI)		(95% CI)	
SpA vs. Controls	415/372	0.052	0.072	0.70 (0.46 to 1.07)	0.095	383	0.042	0.064	0.57 (0.34 to 0.97)	<b>0.028</b>	0.76 (0.56 to 1.02)	0.061
AS vs. Controls	217/372	0.025	0.072	0.33 (0.17 to 0.66)	<b>4.5x10<sup>-4</sup></b>	218	0.026	0.07	0.27 (0.11 to 0.68)	<b>4.0x10<sup>-3</sup></b>	0.35 (0.20 to 0.59)	<b>4.6x10<sup>-6</sup></b>
Non-AS vs. Controls	198/372	0.074	0.072	1.03 (0.63 to 1.68)	0.912	165	0.057	0.056	1.13 (0.57 to 2.27)	0.723	1.18 (0.80 to 1.72)	0.41
AS vs. non-AS	217/198	0.025	0.074	0.32 (0.15 to 0.68)	<b>1.5x10<sup>-3</sup></b>	ND	ND	ND	ND	ND	ND	ND

SpA, spondyloarthritis; AS, ankylosing spondylitis; Non-AS, spondyloarthritis without radiographic sacroiliitis; MAF, minor allelic frequency; T, transmitted minor allelic frequency; U, untransmitted minor allelic frequency; ND, not done

**Table 3.** Results of genotype – phenotype association analysis

	Univariate analysis		Multivariate analysis	
	Median <i>P</i> (IQR) *	Median OR (95% CI)*	Median <i>P</i> (IQR)*	Median OR (95% CI)*
Radiographic SI	<b>0.003</b> (0.001-0.008)	0.44 (0.25 to 0.76)	<b>0.009</b> (0.004-0.02)	0.47 (0.27 to 0.83)
Psoriasis	0.68	0.9 (0.51 to 1.59)	ND	ND
Uveitis	0.36	0.75 (0.41 to 1.38)	ND	ND
IBD	0.24	0.49 (0.15 to 1.61)	ND	ND
Sex	0.24	1.35 (0.82 to 2.23)	ND	ND
Age at onset	<b>0.004</b> (0.002-0.01)	1.03 (1.01 to 1.05)	<b>0.008</b> (0.003-0.02)	1.03 (1.01 to 1.05)

SI: sacro-iliitis; IQR: interquartile range; OR: odds ratio; ND: not done ; IBD: inflammatory bowel disease

\* Median and IQR of the *P*-values and ORs obtained by iterative (10,000. times) random selection of patients

### **Monocyte-Derived Dendritic Cells from HLA-B27+ Axial Spondyloarthritis (SpA) Patients Display Altered Functional Capacity and Dereglated Gene Expression**

*Talpin Alice, Costantino Félicie, Bonilla Nelly, Leboime Ariane, Letourneur Franck, Jacques Sébastien, Dumont Florent, Amraoui Sonia, Dutertre Charles-Antoine, Garchon Henri-Jean, Breban Maxime, Chiocchia Gilles*

Article publié dans « Arthritis Research & Therapy » en août 2014.

### **Les cellules dendritiques dérivées de monocytes de patients HLA-B27+ atteints de spondyloarthrite axiale ont une capacité fonctionnelle altérée et une dérégulation de l'expression de leurs gènes**

De nombreux arguments plaident en faveur de l'implication des cellules dendritiques (DC) dans la pathogénèse de la SpA chez le rat transgénique HLA-B27. En effet, dans ce modèle, des expériences de transfert cellulaire ont suggéré que la présence de cellules présentatrices de l'antigène exprimant fortement le HLA-B27 était nécessaire et suffisante pour provoquer une SpA chez des rats non transgéniques<sup>408,409</sup>. L'étude des DCs de ces rats a également mis en évidence plusieurs anomalies fonctionnelles, telles qu'une diminution de leur capacité à induire une réponse cellulaire T, des anomalies du cytosquelette, une plus faible expression des molécules du CMH de classe II et une augmentation de la mort par apoptose<sup>410-413</sup>. Enfin, ces cellules induisent préférentiellement l'expansion de cellules Th17 proinflammatoires<sup>414</sup>.

En revanche, peu de données sont disponibles quant au rôle des DCs dans la survenue de la SpA chez l'homme. Par analogie à ce qui avait été observé chez le rat, l'objectif de cette étude était de comparer les capacités de stimulation des lymphocytes T CD4+ (LT-CD4+) des DCs dérivées de monocytes du sang circulant (MD-DCs), ainsi que leur profil d'expression génique entre des sujets sains et des sujets atteints de SpA, afin de rechercher des anomalies fonctionnelles des DCs chez ces patients. L'une des doctorantes du laboratoire (Alice Talpin) a réalisé les études cellulaires, tandis que j'ai mené conjointement avec elle les études transcriptomiques (analyse de puce, qPCR, analyse de réseau) ainsi que l'écriture de l'article.

À partir du sang périphérique, les monocytes circulants ont été isolés par sélection CD14+ puis cultivés pendant six jours en présence de GM-CSF et d'IL4. Les MD-DCs ainsi obtenues ont ensuite été laissées au repos ou stimulées par du lipopolysaccharide (LPS) pendant six ou 24 heures.

Dans un premier temps, la capacité des MD-DCs à stimuler une réponse proliférative de LT-CD4+ hétérologues a été testée chez 19 patients et 24 témoins sains. Des LT-CD4+ provenant de témoins sains ont été mis en culture durant six jours en présence des MD-DCs, laissées au repos ou stimulées par le LPS. Une diminution significative de la prolifération des LT-CD4+ par les MD-DCs avant stimulation et après six heures d'exposition au LPS ( $p < 0,05$  et  $p < 0,01$  respectivement) a été observée chez les patients en comparaison aux témoins.

Nous avons ensuite comparé le transcriptome des MD-DCs (laissées au repos ou stimulées par du LPS pendant six ou 24 heures) provenant de neuf patients atteints de SpA axiale avec celui de dix individus sains. Un ensemble de 81 gènes différentiellement exprimés entre patients et témoins (c'est-à-dire avec une valeur  $p$  nominale inférieure à 0,01 et un fold-change inférieur à 0,66 ou supérieur à 1,5) a été identifié. L'analyse d'enrichissement n'a pas permis de mettre en évidence de voie de signalisation spécifiquement dérégulée. Quatre de ces gènes choisis sur la base de leur niveau de significativité statistique et de leur pertinence biologique (*ADAMTS15*, *F13A1*, *SELL* et *CITED2*) ont été validés par PCR quantitative dans les mêmes échantillons.

L'analyse de la co-expression des gènes a permis de montrer qu'il existait une corrélation négative entre l'expression d'*ADAMTS15* et celle de *CITED2* après 6 heures de stimulation des MD-DCs par le LPS. Étant donné son rôle de co-activateur transcriptionnel, nous avons ensuite cherché à identifier les gènes co-régulés avec *CITED2*. On notait, parmi les 222 gènes dont l'expression était significativement corrélée à celle de *CITED2* chez les patients (au seuil de  $p < 0,001$ ), un enrichissement significatif en gènes de la voie Wnt. Cet enrichissement n'existait en revanche pas chez les témoins, soulignant l'implication possible de la voie Wnt dans la physiopathologie de la SpA.

Ce travail a permis d'observer des anomalies fonctionnelles des MD-DCs atteints de SpA de façon similaire à ce qui avait été décrit chez le rat transgénique HLA-B27. La comparaison du transcriptome de MD-DCs entre les patients atteints de SpA et les témoins ne nous a pas permis d'expliquer de façon évidente ces anomalies, mais a mis en évidence l'implication potentielle de certaines voies de signalisation comme la voie Wnt dans la pathogénie de la SpA.



**Monocyte-derived dendritic cells from HLA-B27<sup>+</sup> axial spondyloarthritis (SpA) patients display altered functional capacity and deregulated gene expression**

*Alice Talpin<sup>1,2,3</sup>, Félicie Costantino<sup>1,2,3,4</sup>, Nelly Bonilla<sup>5</sup>, Ariane Leboime<sup>4</sup>, Franck Letourneur<sup>5</sup>, Sébastien Jacques<sup>5</sup>, Florent Dumont<sup>5</sup>, Sonia Amraoui<sup>5,6</sup>, Charles-Antoine Dutertre<sup>7</sup>, Henri-Jean Garchon<sup>1,2,3,8</sup>, Maxime Breban<sup>1,2,3,4</sup>, Gilles Chiochia<sup>1,2,3\*</sup>*

*1 – INSERM U987, Université Versailles-Saint-Quentin, Versailles, 78000, France;*

*2 – Laboratoire d'excellence INFLAMEX, Sorbonne Paris-Cité, Paris, France;*

*3 – UFR des Sciences de la Santé, Simone Veil, Versailles Saint Quentin en Yvelines Université, Montigny-Le-Bretonneux, 78180, France;*

*4 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, 92100, France;*

*5 – Institut Cochin, INSERM U1016, CNRS (UMR 8104), Université Paris-Descartes, Sorbonne Paris-Cité, 75014, France;*

*6 – Antigen Presentation by Dendritic Cell Team; Institut Cochin, INSERM U1016, CNRS (UMR 8104), Université Paris-Descartes, Sorbonne Paris-Cité, 75014, France;*

*7 – Duke-NUS Graduate Medical School, Program in Emerging Infectious Disease, Singapore, Singapore;*

*8 – Genetics Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, 92100, France;*

\* Corresponding author

Email: [gilles.chiochia@inserm.fr](mailto:gilles.chiochia@inserm.fr)

**Keywords:** spondyloarthritis, ankylosing spondylitis, transcriptomic, monocyte-derived dendritic cells

## ABSTRACT

**Objectives:** To compare the functional capacity and gene expression profile of monocyte-derived dendritic cells (MD-DCs) in HLA-B27<sup>+</sup> axial spondyloarthritis (SpA) patients and healthy controls.

**Methods:** MD-DCs were differentiated with IL-4 and GM-CSF for 7 days, starting from purified CD14<sup>+</sup> monocytes and stimulated with LPS for 6 and 24 hours. Their capacity to stimulate allogeneic CD4<sup>+</sup> T cells from unrelated healthy donor was tested. Transcriptomic study was performed with Affymetrix HuGene 1.0 ST microarrays. Gene expression levels were compared between patients and controls using a multivariate design under a linear model (LIMMA). Real-time quantitative PCR (qRT-PCR) was performed for validation of the most striking gene expression differences.

**Results:** The stimulatory capacity of allogeneic CD4<sup>+</sup> T cells by MD-DCs from SpA patients was decreased. Transcriptomic analysis revealed 81 genes differentially expressed in MD-DCs between SpA patients and controls ( $p < 0.01$  and fold-change  $< 0.66$  or  $> 1.5$ ). Four selected genes were validated by qRT-PCR: *ADAMTS15*, *CITED2*, *F13A1* and *SELL*. Expression levels of *ADAMTS15* and *CITED2*, encoding a metallopeptidase and a transcription factor, respectively, were inversely correlated with each other ( $R=0.75$ ,  $p=0.0003$ ). Furthermore, *in silico* analysis identified several genes of the Wnt signaling pathway having expression co-regulated with *CITED2*.

**Conclusion:** This study revealed altered function and gene expression pattern in MD-DCs from HLA-B27<sup>+</sup> axial SpA. Co-expression study showed an inverse correlation between *ADAMTS15* and *CITED2*. Moreover, the Wnt signaling pathway appeared as deregulated in SpA MD-DCs, a finding which may be connected to Th17-driven inflammatory responses.

## INTRODUCTION

Spondyloarthritis (SpA) is a chronic inflammatory rheumatic disorder, with a prevalence around 0.42% in Caucasian population [1]. Joint inflammation is responsible for pain and stiffness but long-term outcome is mainly determined by new bone formation, which can lead to complete ankylosis. Effective treatments such as TNF blockers can temporarily suppress inflammation, but none has yet proven to affect long-term disease outcome. Therefore, there is a need to better understand pathological mechanisms controlling both initiation and progression of SpA.

Genome-wide gene expression analysis is a powerful approach to identify molecular mechanisms responsible for a disease. Microarray studies have already been conducted in SpA, most of them focusing on the ankylosing spondylitis (AS) subtype [2-7]. Each of those studies identified several genes differentially expressed between patients and controls but there was very little overlap between their results, maybe due to the variety of protocols.

Several studies support an important role of dendritic cells (DCs) in the pathogenesis of SpA. Indeed, aberrant functions of DCs have been demonstrated in an HLA-B27/human  $\beta$ 2-microglobulin transgenic rat model of SpA [8]. First, DCs from these rats have a decreased capacity to stimulate primary allogeneic or syngeneic T cell responses [9]. Furthermore, the proportion of conjugates formed between HLA-B27 DCs and naïve CD4<sup>+</sup> T cells is reduced. Moreover, mature HLA-B27 molecules expressed by DCs appear to impair the formation of an antigen-independent immunologic synapse with naïve CD4<sup>+</sup> T cells by interfering with the engagement of costimulatory molecules [10]. Finally, B27 transgenic rat DCs were shown to favor Th17 expansion [11,12] and to alter regulatory T cells function, resulting in decreased IL-10 and enhanced IL-17 production [13].

On this basis, we hypothesized that DCs also play an important role in human disease, as proposed in the HLA-B27 transgenic rat model. The aim of the current study was to compare functional capacity and gene expression profile of monocyte-derived DCs (MD-DCs) in HLA-B27<sup>+</sup> axial SpA patients and healthy controls.

## PATIENTS AND METHODS

### Patients and controls

We studied three different groups of patients and controls: 10 patients and 5 controls for monocyte subsets characterization by flow cytometry, 19 patients and 24 controls for mixed

lymphocyte reaction and 9 patients and 10 controls for the MD-DCs transcriptomic study (6 patients and 4 controls were investigated for both the latters). All the patients were HLA-B27<sup>+</sup> and fulfilled the Assessment of SpondyloArthritis international Society classification criteria for axial SpA [14]. Controls were healthy blood donors.

Patients and healthy controls characteristics are summarized in additional files (see **Additional file 1: Table S1** and **Additional file 2: Table S2**, respectively). All participants to the study gave written informed consent and the study was approved by local ethics committee of Ile-de-France XI (Saint-Germaine-en-Laye France).

### **Cell isolation, culture and stimulation**

Peripheral blood mononuclear cells (PBMCs) were isolated from 50 mL of blood by gradient separation on Ficoll density gradient centrifugation (STEMCELL Technologies).

Monocytes used to generate MD-DCs were purified by magnetic cell sorting using anti-CD14 monoclonal antibody (mAb)-coated beads (BD IMag). Sorted monocytes were morphologically homogeneous with 99% of CD14<sup>+</sup> cells, as determined by flow cytometry.

Monocytes were further cultured for 6 days in 24-well plates (400,000 cells/500  $\mu$ L) in RPMI 1640 medium supplemented with 10% heat inactivated fetal calf serum, 100 U/mL penicillin, 100  $\mu$ g/mL streptomycin, 500 U/mL recombinant human granulocyte-macrophage colony stimulating factor (rhGM-CSF) and 500 U/mL rhIL-4 (AbCys S.A., France). Then, the MD-DCs were stimulated or not with lipopolysaccharide from E. Coli (LPS, Sigma-Aldrich, St. Louis, MO) at a concentration of 100 ng/mL for the last 6 or 24 hours of culture (further referred to as time-points H0, H6 and H24).

CD4<sup>+</sup> T cells were purified from two unrelated healthy donor PBMCs by magnetic cell sorting using anti-CD4 monoclonal antibody (mAb)-coated beads (BD IMag), and stored frozen until being used for mixed lymphocyte reaction (MLR).

### **Flow cytometry**

To characterize monocyte subsets, freshly purified PBMCs were analyzed by 6-color flow cytometry on a FACS LSRII apparatus. The gating strategy was based on a previous report [15]. Monocytes were subdivided into three major subsets: classical CD14<sup>++</sup>CD16<sup>-</sup>, intermediate CD14<sup>++</sup>CD16<sup>+</sup> and non-classical CD14<sup>+</sup>CD16<sup>++</sup> monocytes. The following anti-human mAbs were used: CD45-Amcyan (BD Biosciences), HLA-DR-PerCP (BD

Biosciences), CD19-ECD (Beckman Coulter), CD14-QDot655 (Invitrogen), CD16-APC-H7 (BD Biosciences). The Live/Dead blue Dye (Invitrogen) was used to exclude dead cells.

Samples of the purified monocytes used to generate MD-DCs and of the resulting MD-DCs were routinely stained with the following anti-human mAbs: CD14-FITC, CD11c-APC, CD40-PE, HLA-I-FITC, HLA-DR-PerCP, CD80-PE, CD83-APC and CD86-FITC (all from BD Bioscience) and analyzed by flow cytometry on a FACS canto II apparatus (BD Biosciences).

### **Mixed Lymphocyte Reaction (MLR)**

Purified allogeneic CD4<sup>+</sup> T cells (10<sup>5</sup> cells per well) from healthy donor were cultured with unstimulated (H0) or LPS-stimulated (H6, H24) MD-DCs (10<sup>4</sup> cells per well), in 96-well flat-bottomed culture dishes in a final volume of 200 µL. Proliferation of T cells was assayed by measuring incorporation of <sup>3</sup>H-deoxythymidine added (0.5 µCi per well) after 6 days of culture, using a Microbeta scintillation counter (Wallac, Turku, Finland). Data are expressed as the mean counts per minute (CPM) in triplicate wells. An MLR index (ratio of CPM of MLR on CPM of CD4<sup>+</sup> T cells only) was used to represent CD4<sup>+</sup> T cells proliferation. Two CD4<sup>+</sup> T cell batches stored from different healthy donors were sequentially used for MLR in two sets of experiments, each including equivalent numbers of patient and control MD-DC samples. Since there was no statistically significant difference in the results between both sets of experiments, we pooled them.

Wilcoxon test was used to compare MLR indices between patients and controls at each stimulation time-point.

### **Transcriptomic study**

#### *RNA isolation*

MD-DCs were disrupted and homogenized using RLT buffer (Qiagen, Valencia, CA). Total RNA was isolated using RNeasy Mini Kit (Qiagen, Valencia, CA). RNA quantity and quality were assessed using Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Only samples with an RNA integrity number (RIN) above 8 were further processed.

### *Microarray hybridization*

RNA was reverse transcribed, converted to biotinylated complementary RNA using standard Affymetrix protocol (Affymetrix, Santa Clara, CA) and hybridized to the Affymetrix GeneChip Human Gene 1.0 ST Array by the genomic platform of Cochin Institute.

### *Differential gene expression validation by qRT-PCR*

For validation, the relative gene expression levels of candidate genes identified through the foregoing microarray study were further quantified using qRT-PCR. Briefly, RNA treated with DNase I (Invitrogen) was reverse transcribed using SuperscriptII (Invitrogen) and then quantified using the SYBR green PCR Master Mix (Applied Biosystems) and the 7300 Real-Time PCR System (Applied Biosystems). Primers were purchased from Eurofins MWG (nucleotide sequences of the PCR primers are available in **Additional file 3: Table S3**). The experiment design included three technical replicates.

### *Statistical analysis*

Raw Affymetrix data (.cel files) from 57 arrays (corresponding to H0, H6 and H24 time-points of stimulation by LPS for 19 subjects) were transformed by Robust Multichip Analysis (RMA) method using Bioconductor in R software (library Affy) [16]. This transformation included background correction, normalization and summarization of expression values using Brainarray version 15 custom chip definition (cdf) files generated with the Ensembl annotation set [17]. Before analysis, genes were filtered on annotation and expression level mean with a cut-off at 6.11 (corresponding to the mean of expression level of both anti-genomic and intronic probes of housekeeping genes). Gene expression levels were then fitted to a bivariate linear model including disease status and time-point after stimulation, using the Linear Models for Microarray data (LIMMA) package of Bioconductor in R environment to determine differentially expressed genes [18]. The threshold for differential gene expression between patients and controls was set to a global fold change of more or less than 1.5 with a nominal p-value less than 0.01.

For qRT-PCR validation, gene expression data were computed with the  $\Delta\Delta C_q$  method, using three housekeeping genes: *RPL30*,  *$\beta$ -ACT* and *GAPDH*. Then, 2-way analysis of variance (ANOVA) was used to test differential gene expression among samples from SpA and controls without (baseline; no treatment) and with LPS treatment for 6 and 24 hours. In

case of statistical significance ( $p < 0.05$ ), post-hoc  $t$  tests with Bonferroni correction allowed to perform pairwise comparisons.

### **Paired genes co-expression study and molecular pathway analysis**

Pairwise correlations between gene expression levels were tested using the nonparametric Spearman test. We used the commercial software Genomatix Pathway System (<http://www.genomatix.de>) to identify significantly enriched pathways and functional themes, among the list of the best correlated genes with genes of interest.

## **RESULTS**

### **Functional impairment of MD-DCs from SpA patients**

To verify that MD-DCs were differentiated from comparable monocyte populations between SpA patients and controls, we analyzed monocytes subsets in both groups. As shown in an additional file, the distribution of the classical, intermediate and non-classical circulating subsets among CD45<sup>+</sup> PBMCs was similar in patients and controls (see **Additional file 4: Figure S4**).

The cells obtained after 7 days of differentiation in culture in patients and controls were uniformly CD14<sup>-</sup>, CD11c<sup>+</sup>, CD40<sup>+</sup>, HLA-DR<sup>+</sup>, CD83<sup>dim</sup> and CD86<sup>+</sup>, consistent with a MD-DC phenotype. Moreover, HLA-DR, CD80 and CD86 expression increased and CD83 was induced on MD-DCs after LPS stimulation for 24 hours (**Additional file 5: Figure S5**).

In rat model of SpA, HLA-B27 transgenic DCs have a decreased capacity to stimulate allogeneic T cells. Thus, we tested the capacity of MD-DCs from HLA-B27<sup>+</sup> SpA patients to prime allogeneic CD4<sup>+</sup> T cells using an *in vitro* proliferation assay. A weaker proliferation of CD4<sup>+</sup> T cells was observed with SpA MD-DCs as compared to controls, that was statistically significant before and after 6 hours of LPS exposure ( $p < 0.05$  and  $p < 0.01$ , respectively; **Figure 1**).

### **Identification of differentially expressed genes in MD-DCs between patients and controls**

To investigate the mechanisms underlying the functional defect of MD-DCs from SpA patients, we then compared gene expression levels in those cells between SpA and healthy donors upon stimulation with LPS.

After filtering on gene expression level and gene annotation, 13,021 genes (57.3% of the genes present in the microarray) were kept for further analysis. Expression data for MD-DCs either left unstimulated or stimulated with LPS for 6 or 24 hours were first globally compared between SpA patients and healthy subjects. Unsupervised hierarchical clustering showed that samples were perfectly grouped according to LPS stimulation time-points (**Additional file 6: Figure S6**) but not to the disease status. Linear modeling nevertheless identified 81 genes differentially expressed between patients and controls at any time-point, with a nominal p-value <0.01 and a fold-change lesser than 0.66 or greater than 1.5 (**Additional file 7: Table S7**). Of this set of genes, 61 were down-regulated and 20 were up-regulated in patients.

Using the extraction of the dataset corresponding to these 81 genes, LPS stimulation time-points were perfectly separated by unsupervised hierarchical clustering. Furthermore, patients and controls were now clearly discriminated, with only four misclassifications (one at H0, two at H6 and one at H24) (**Figure 2**).

Four candidate genes were chosen for validation study on the basis of their p-value, fold-change and biological relevance: *ADAMTS15*, *CITED2*, *F13A1* and *SELL*. For all 4 genes, the qRT-PCR data produced with the RNA samples used for the microarray study confirmed the significant differences in expression and the direction of changes between SpA patients and controls (**Table 1, Figure 3**): *ADAMTS15*, *F13A1* and *SELL* were significantly up-regulated in SpA samples whereas *CITED2* was down-regulated.

### Co-regulated genes in SpA patients

Examination of gene expression data indicated that some of them displayed similarity in their expression profile suggesting co-regulation. Indeed, we found a highly significant negative correlation between *ADAMTS15* and *CITED2* expression levels at H6 in both patient and control groups (combined data:  $\rho = -0.75$ ;  $P = 0.0003$ ; **Figure 4**).

Considering the transcriptional coactivator function of *CITED2* and its down-regulation in SpA MD-DCs, we sought whether other genes than *ADAMTS15* shared a similarly co-regulated expression pattern in patients MD-DCs. We tested the correlation of genome-wide expression matrix of patients MD-DCs with *CITED2* expression in a pairwise fashion. A set of 222 genes was found significantly correlated with *CITED2* ( $p < 0.01$ ). They were analyzed with the Genomatix web platform in order to interrogate their biological



relevance. This revealed a significant enrichment of genes belonging to the Wnt signaling pathway ( $p= 2.48 \times 10^{-4}$ ) in the SpA group. In the controls, there was no significant correlation between expression levels of *CITED2* and the four genes identified in the Wnt pathway (i.e. *WNT1*, *WNT10B*, *FZD4* and *ROR2*) (**Figure 5**). This co-regulation pattern involving four genes of the Wnt signaling pathway was therefore specifically associated with the disease.

## DISCUSSION

Starting from the hypothesis that aberrant DCs function could play a critical role in the development of SpA, we showed here for the first time that the capacity of MD-DCs to stimulate allogeneic CD4<sup>+</sup> T cell response was impaired in SpA patients compared to controls. The genome-wide transcriptome of these MD-DCs elicited by LPS stimulation revealed 81 genes differentially expressed of whom 4 have been validated by RT-qPCR. Finally, we identified the Wnt signaling pathway as dysregulated in patients.

Our choice to study *in vitro*-differentiated DCs rather than DCs purified from peripheral blood had the advantage to remove those cells from *in vivo* influences, such as non-specific acute inflammation and drug therapy. Therefore, the differences that we observed between DCs from SpA and controls were more likely to be intrinsic to the DC and could be involved in the disease mechanism.

First, we demonstrated that MD-DCs differentiated from similar monocyte populations in both groups. Then, we observed that the ability of SpA MD-DCs to stimulate allogeneic CD4<sup>+</sup> T cell proliferation was markedly defective, as compared to controls, as previously shown in SpA-prone HLA-B27 transgenic rats. The molecular basis for this impaired DCs function has yet to be elucidated.

To investigate this impairment, we have studied the time-dependent global transcriptome of MD-DCs in response to LPS stimulation. A number of microarray-based studies have previously been undertaken in AS, on whole blood cells [3, 5], PBMCs [6, 7] or macrophages [2]. But no transcriptomic analysis had been conducted on DCs. Of note, as we worked on a homogeneous cell population rather than on cell mixture, our results were not confounded by variations in the composition of the cell populations [19]. The time-dependency of our gene profiling is also a novel aspect, rarely investigated in humans.

Our analysis revealed 81 genes differentially expressed in resting and/or LPS-stimulated MD-DCs between SpA patients and controls. Of note, we did not identify a reverse interferon signature, such as in HLA-B27 transgenic rat splenic DCs [20] or in SpA patients monocyte-derived macrophages [2]. Such discrepancy could be explained by differences in culture conditions. Here, we used IL-4 to differentiate MD-DCs, a cytokine that opposes interferon and interferon-induced genes expression and this might have blunted interferon-related differences between groups. Based on the magnitude of their variation between patients and controls, the level of statistical significance, and their biological relevance, we selected four of these genes that we validated using qRT-PCR.

We found an increased expression of *ADAMTS15* (A Disintegrin And Metalloproteinase with ThromboSpondin motif) in SpA. Although ADAMTS family has not yet been implicated in this disorder, numerous studies have identified a role for metalloproteinases (MMPs) in SpA susceptibility and severity [21-23]. Noteworthy, *ADAMTS15* was shown to be expressed in the joint with decreased expression in osteoarthritis [24]. Thus, *ADAMTS15* could be implicated in the cartilage and/or bone turn-over that takes place during joint inflammation, such as in SpA.

*CITED2* functions as a context-dependent transcriptional modulator to up- or down-regulate the expression of specific genes [25,27]. Here, the expression of *CITED2* was down-regulated in SpA MD-DCs. We further showed an inverse correlation between *CITED2* and *ADAMTS15* expression after six hours of LPS treatment. Interestingly, similar inverse correlation has previously been reported between *CITED2* and several MMP family members [28, 29]. MMPs constitute a very important group of proteolytic enzymes in joint tissues. Thus, it has been suggested that *CITED2* exerted chondroprotective effects through MMPs down-regulation [30]. The down-regulation of *CITED2* and the inverse correlation between *CITED2* and *ADAMTS15* expressions that we observed here suggest that DCs or other phagocytic mononuclear cells, such as macrophages and osteoclasts, could be implicated in SpA joint resorption through heightened metallopeptidases activity.

*In silico* pathway analysis conducted on *CITED2* co-expressed genes highlighted a down-regulation of several factors belonging to the canonical (i.e. *WNT1*, *WNT10B* and *FRZL4*) and non-canonical (*ROR2*) Wnt signaling pathways, of potential relevance for SpA pathogenesis. First, factors of the canonical Wnt pathway were shown to play a crucial role either in bone formation or destruction in inflammatory arthritis [31, 32], such as *WNT10b*

[33, 34]. On the other hand, activation of canonical Wnt- $\beta$ -catenin signaling in DCs was shown to concur to regulatory T cells differentiation and conversely to inhibit Th17 differentiation [35]. Moreover, ROR2 is a signaling component of the non-canonical Wnt pathway acting on actin cytoskeleton to stimulate cell migration [36, 37]. Interestingly, altered T cell stimulation was linked to defective cytoskeleton dynamics in HLA-B27 transgenic rat DCs [38].

*F13A1* encodes the coagulation factor XIII A subunit, a transglutaminase enzyme. It has multiple extra- and intra-cellular functions, including a role in cartilage and bone development. Hence, increased factor XIII A expression has previously been associated with cartilage aging and degeneration [39]. Fibrin crosslinking by factor XIII is of crucial importance not only for haemostasis, but also for inflammation. For instance, factor XIII A-subunit genotype was shown to influence C-reactive protein levels during inflammation in rheumatoid arthritis (RA) [40].

Finally, *SELL* encodes the lymphocyte homing receptor L-selectin/CD62L, one of the major adhesion molecules, which regulates entry of neutrophils and monocytes into inflamed tissues and contributes to the severity of joint inflammation in experimental arthritis [41]. Thus, up-regulation of *SELL*, as shown here in DCs could well participate to joint inflammation in SpA. Interestingly, this gene was also found upregulated in splenic DCs from HLA-B27/human  $\beta$ 2-microglobulin transgenic rat [20].

## CONCLUSION

Results of the present study reveal the defective functional capacity of DCs from SpA patients, as compared to controls. Furthermore, our results demonstrate significant changes in MD-DCs gene expression upon LPS stimulation that may be inherent to SpA patients. Some of them (i.e. decreased Wnt signaling) could account for the altered DCs function that provided a rationale for the present study. Others (i.e. up-regulation of *ADAMTS15*, *F13A1* and *SELL*) would concur to reinforce tissue inflammation and/or damage. Finally, the co-expression of *CITED2* co-transcriptional factor with several of the foregoing genes (i.e. Wnt signaling pathway and *ADAMTS15*) supports the hypothesis that a coordinate deregulation taking place in DCs may play an important role in SpA pathogenesis.

## REFERENCES

1. Costantino F, Talpin A, Said-Nahal R, Goldberg M, Henny J, Chiocchia G, Garchon H-J, Zins M, Breban M: Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann Rheum Dis* 2013. Dec 18. doi: 10.1136/annrheumdis-2013-204436. [Epub ahead of print]
2. Smith JA, Barnes MD, Hong D, DeLay ML, Inman RD, Colbert RA: Gene expression analysis of macrophages derived from ankylosing spondylitis patients reveals interferon-gamma dysregulation. *Arthritis Rheum* 2008, 58:1640–1649.
3. Gu J, Wei Y-L, Wei JCC, Huang F, Jan M-S, Centola M, Frank MB, Yu D: Identification of RGS1 as a candidate biomarker for undifferentiated spondylarthritis by genome-wide expression profiling and real-time polymerase chain reaction. *Arthritis Rheum* 2009, 60:3269–3279.
4. Duan R, Leo P, Bradbury L, Brown MA, Thomas G: Gene expression profiling reveals a downregulation in immune-associated genes in patients with AS. *Ann Rheum Dis* 2010, 69:1724–1729.
5. Assassi S, Reveille JD, Arnett FC, Weisman MH, Ward MM, Agarwal SK, Gourh P, Bhula J, Sharif R, Sampat K, Mayes MD, Tan FK: Whole-blood gene expression profiling in ankylosing spondylitis shows upregulation of toll-like receptor 4 and 5. *J Rheumatol* 2011, 38:87–98.
6. Pimentel-Santos FM, Ligeiro D, Matos M, Mourão AF, Costa J, Santos H, Barcelos A, Godinho F, Pinto P, Cruz M, Fonseca JE, Guedes-Pinto H, Branco JC, Brown MA, Thomas GP: Whole blood transcriptional profiling in ankylosing spondylitis identifies novel candidate genes that might contribute to the inflammatory and tissue-destructive disease aspects. *Arthritis Res Ther* 2011, 13:R57.
7. Sharma SM, Choi D, Planck SR, Harrington CA, Austin CR, Lewis JA, Diebel TN, Martin TM, Smith JR, Rosenbaum JT: Insights in to the pathogenesis of axial spondyloarthritis based on gene expression profiles. *Arthritis Res Ther* 2009, 11:R168.
8. Breban M, Hacquard-Bouder C, Falgarone G: Animal models of HLA-B27-associated diseases. *Curr Mol Med* 2004, 4:31–40.
9. Hacquard-Bouder C, Falgarone G, Bosquet A, Smaoui F, Monnet D, Ittah M, Breban M: Defective costimulatory function is a striking feature of antigen-presenting cells in an HLA-B27-transgenic rat model of spondylarthropathy. *Arthritis Rheum* 2004, 50:1624–1635.

10. Hacquard-Bouder C, Chimenti M-S, Giquel B, Donnadieu E, Fert I, Schmitt A, André C, Breban M: Alteration of antigen-independent immunologic synapse formation between dendritic cells from HLA-B27-transgenic rats and CD4<sup>+</sup> T cells: selective impairment of costimulatory molecule engagement by mature HLA-B27. *Arthritis Rheum* 2007, 56:1478–1489.
11. Glatigny S, Fert I, Blaton MA, Lories RJ, Araujo LM, Chiocchia G, Breban M: Proinflammatory Th17 cells are expanded and induced by dendritic cells in spondylarthritis-prone HLA-B27-transgenic rats. *Arthritis Rheum* 2012, 64:110–120.
12. Utriainen L, Firmin D, Wright P, Cerovic V, Breban M, McInnes I, Milling S: Expression of HLA-B27 causes loss of migratory dendritic cells in a rat model of spondylarthritis. *Arthritis Rheum* 2012, 64:3199–3209.
13. Araujo L, Fert I, Jouhault Q, Labroquère K, Andrieu M, Chiocchia G, Breban M: Increased production of interleukin-17 over interleukin-10 by regulatory T cells implicates ICOS molecule in experimental spondyloarthritis. *Arthritis Rheumatol* 2014.
14. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, Braun J, Chou CT, Collantes-Estevez E, Dougados M, Huang F, Gu J, Khan MA, Kirazli Y, Maksymowych WP, Mielants H, Sørensen IJ, Ozgocmen S, Roussou E, Valle-Oñate R, Weber U, Wei J, Sieper J: The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009, 68:777–783.
15. Dutertre C-A, Amraoui S, DeRosa A, Jourdain J-P, Vimeux L, Goguet M, Degrelle S, Feuillet V, Liovat A-S, Müller-Trutwin M, Decroix N, Deveau C, Meyer L, Goujard C, Loulergue P, Launay O, Richard Y, Hosmalin A: Pivotal role of M-DC8<sup>+</sup> monocytes from viremic HIV-infected patients in TNF $\alpha$  overproduction in response to microbial products. *Blood* 2012, 120:2259–2268.
16. Gautier L, Cope L, Bolstad BM, Irizarry RA: affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004, 20:307–315.
17. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F: Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucl Acids Res* 2005, 33:e175–e175.
18. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004, 3:Article3.

19. Plessy C, Desbois L, Fujii T, Carninci P: Population transcriptomics with single-cell resolution: A new field made possible by microfluidics: A technology for high throughput transcript counting and data-driven definition of cell types. *Bioessays* 2012.
20. Fert I, Cagnard N, Glatigny S, Letourneur F, Jacques S, Smith JA, Colbert RA, Taurog JD, Chiocchia G, Araujo LM, Breban M: Reverse interferon signature is characteristic of antigen-presenting cells in human and rat spondyloarthritis. *Arthritis Rheumatol* 2014, 66:841–851.
21. Zhu J, Yu DTY: Matrix metalloproteinase expression in the spondyloarthropathies. *Curr Opin Rheumatol* 2006, 18:364–368.
22. Maksymowych WP, Landewé R, Conner-Spady B, Dougados M, Mielants H, van der Tempel H, Poole AR, Wang N, van der Heijde D: Serum matrix metalloproteinase 3 is an independent predictor of structural damage progression in patients with ankylosing spondylitis. *Arthritis Rheum* 2007, 56:1846–1853.
23. Mathey DL, Packham JC, Nixon NB, Coates L, Creamer P, Hailwood S, Taylor GJ, Bhalla AK: Association of cytokine and matrix metalloproteinase profiles with disease activity and function in ankylosing spondylitis. *Arthritis Res Ther* 2012, 14:R127.
24. Kevorkian L, Young DA, Darrah C, Donell ST, Shepstone L, Porter S, Brockbank SMV, Edwards DR, Parker AE, Clark IM: Expression profiling of metalloproteinases and their inhibitors in cartilage. *Arthritis Rheum* 2004, 50:131–141.
25. Bragança J, Swingler T, Marques FIR, Jones T, Eloranta JJ, Hurst HC, Shioda T, Bhattacharya S: Human CREB-binding protein/p300-interacting transactivator with ED-rich tail (CITED) 4, a new member of the CITED family, functions as a co-activator for transcription factor AP-2. *J Biol Chem* 2002, 277:8559–8565.
26. Lou X, Sun S, Chen W, Zhou Y, Huang Y, Liu X, Shan Y, Wang C: Negative feedback regulation of NF- $\kappa$ B action by CITED2 in the nucleus. *J Immunol* 2011, 186:539–548.
27. Goodman RH, Smolik S: CBP/p300 in cell growth, transformation, and development. *Genes Dev* 2000, 14:1553–1577.
28. Bai L, Merchant JL: A role for CITED2, a CBP/p300 interacting protein, in colon cancer cell invasion. *FEBS Lett* 2007, 581:5904–5910.
29. Lee JY, Taub PJ, Wang L, Clark A, Zhu LL, Maharam ER, Leong DJ, Ramcharan M, Li Z, Liu Z, Ma Y-Z, Sun L, Zaidi M, Majeska RJ, Sun HB: Identification of CITED2 as a negative regulator of fracture healing. *Biochem Biophys Res Commun* 2009, 387:641–645.

30. Leong DJ, Li YH, Gu XI, Sun L, Zhou Z, Nasser P, Laudier DM, Iqbal J, Majeska RJ, Schaffler MB, Goldring MB, Cardoso L, Zaidi M, Sun HB: Physiological loading of joints prevents cartilage degradation through CITED2. *FASEB J* 2011, 25:182–191.
31. Goldring SR, Goldring MB: Eating bone or adding it: the Wnt pathway decides. *Nat Med* 2007, 13:133–134.
32. Laine CM, Joeng KS, Campeau PM, Kiviranta R, Tarkkonen K, Grover M, Lu JT, Pekkinen M, Wessman M, Heino TJ, Nieminen-Pihala V, Aronen M, Laine T, Kröger H, Cole WG, Lehesjoki A-E, Nevarez L, Krakow D, Curry CJR, Cohn DH, Gibbs RA, Lee BH, Mäkitie O: WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. *N Engl J Med* 2013, 368:1809–1816.
33. Bennett CN, Longo KA, Wright WS, Suva LJ, Lane TF, Hankenson KD, MacDougald OA: Regulation of osteoblastogenesis and bone mass by Wnt10b. *Proc Natl Acad Sci USA* 2005, 102:3324–3329.
34. Franck H, Meurer T, Hofbauer LC: Evaluation of bone mineral density, hormones, biochemical markers of bone metabolism, and osteoprotegerin serum levels in patients with ankylosing spondylitis. *J Rheumatol* 2004, 31:2236–2241.
35. Manicassamy S, Reizis B, Ravindran R, Nakaya H, Salazar-Gonzalez RM, Wang Y-C, Pulendran B: Activation of beta-catenin in dendritic cells regulates immunity versus tolerance in the intestine. *Science* 2010, 329:849–853.
36. Nishita M, Yoo SK, Nomachi A, Kani S, Sougawa N, Ohta Y, Takada S, Kikuchi A, Minami Y: Filopodia formation mediated by receptor tyrosine kinase Ror2 is required for Wnt5a-induced cell migration. *J Cell Biol* 2006, 175:555–562.
37. Nomachi A, Nishita M, Inaba D, Enomoto M, Hamasaki M, Minami Y: Receptor tyrosine kinase Ror2 mediates Wnt5a-induced polarized cell migration by activating c-Jun N-terminal kinase via actin-binding protein filamin A. *J Biol Chem* 2008, 283:27973–27981.
38. Dhaenens M, Fert I, Glatigny S, Haerinck S, Poulain C, Donnadieu E, Hacquard-Bouder C, André C, Elewaut D, Deforce D, Breban M: Dendritic cells from spondylarthritis-prone HLA-B27-transgenic rats display altered cytoskeletal dynamics, class II major histocompatibility complex expression, and viability. *Arthritis Rheum* 2009, 60:2622–2632.
39. Johnson K, Hashimoto S, Lotz M, Pritzker K, Terkeltaub R: Interleukin-1 induces pro-mineralizing activity of cartilage tissue transglutaminase and factor XIIIa. *Am J Pathol* 2001, 159:149–163.
40. Hoppe B, Häupl T, Skapenko A, Ziemer S, Tauber R, Salama A, Schulze-Koops H, Burmester G-R, Dörner T: Fibrinogen and factor XIII A-subunit genotypes interactively

influence C-reactive protein levels during inflammation. *Ann Rheum Dis* 2012, 71:1163–1169.

41. Sarraj B, Ludányi K, Glant TT, Finnegan A, Mikecz K: Expression of CD44 and L-selectin in the innate immune system is required for severe joint inflammation in the proteoglycan-induced murine model of rheumatoid arthritis. *J Immunol* 2006, 177:1932–1940.

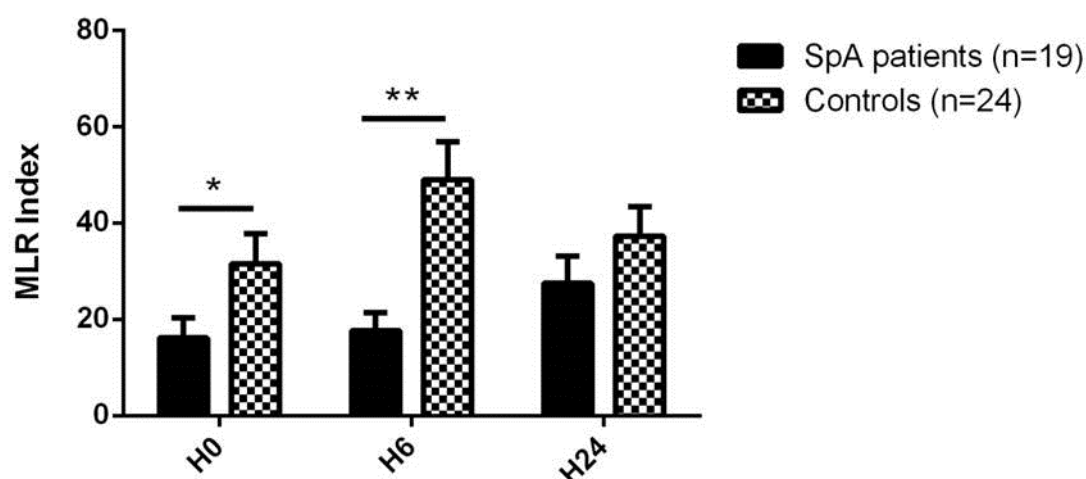


## TABLES

**Table 1. qRT-PCR expression analysis of four selected genes in MD-DCs from SpA patients, as compared to controls.** P-value is reported considering the global effect, the effect of the disease status only or the effect of the disease status adjusted for time.

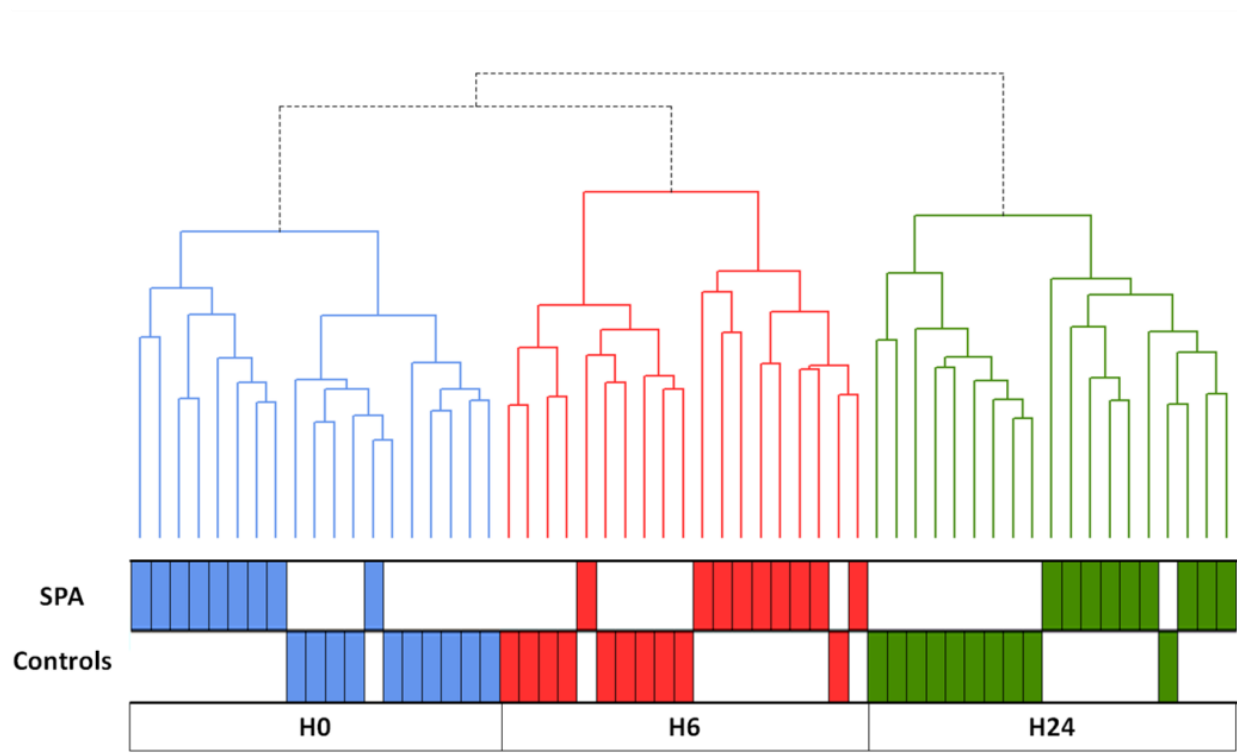
Gene	Microarray	qRT-PCR				
	P-value	P-value		Fold-change		
		Disease	Disease/time	H0	H6	H24
<i>ADAMTS15</i>	0.0003	<b>0.0005</b>	<b>0.04</b>	2.2	<b>5.63</b>	2.2
<i>CITED2</i>	0.0007	0.21	<b>0.031</b>	0.94	<b>0.51</b>	0.99
<i>F13A1</i>	0.00006	<b>0.019</b>	<b>0.001</b>	1.32	1.25	<b>5.05</b>
<i>SELL</i>	0.006	<b>0.012</b>	0.8	<b>3.57</b>	<b>3.47</b>	2.97

## FIGURES

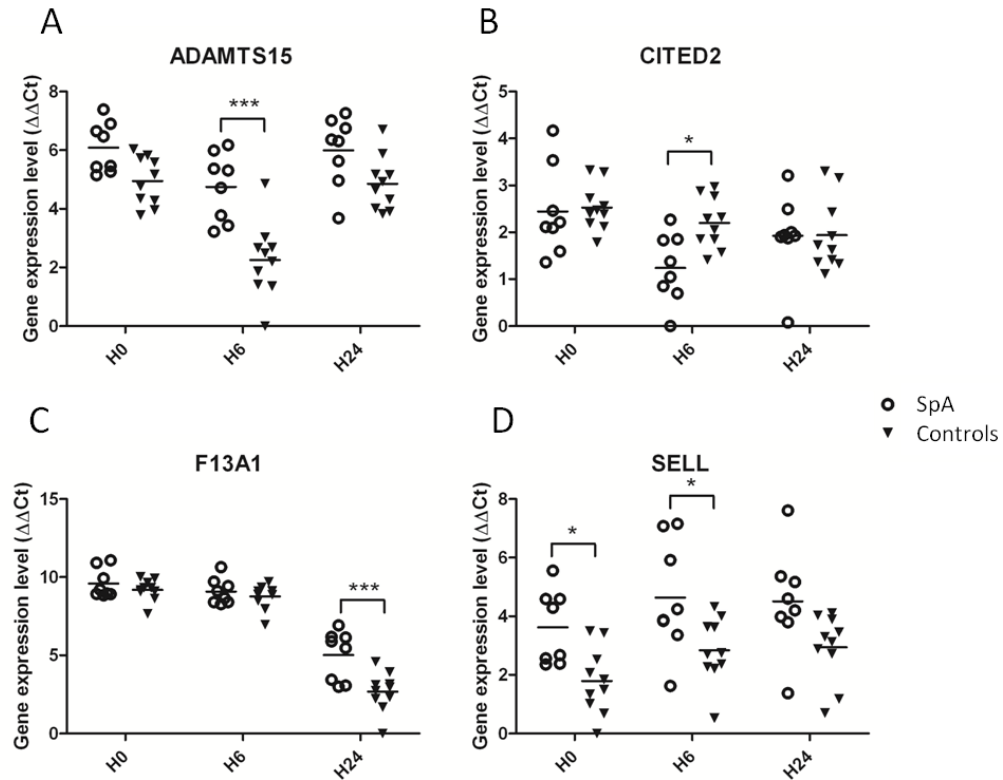


**Figure 1. Impaired allogeneic T cell stimulatory capacity of MD-DCs from SpA patients.**

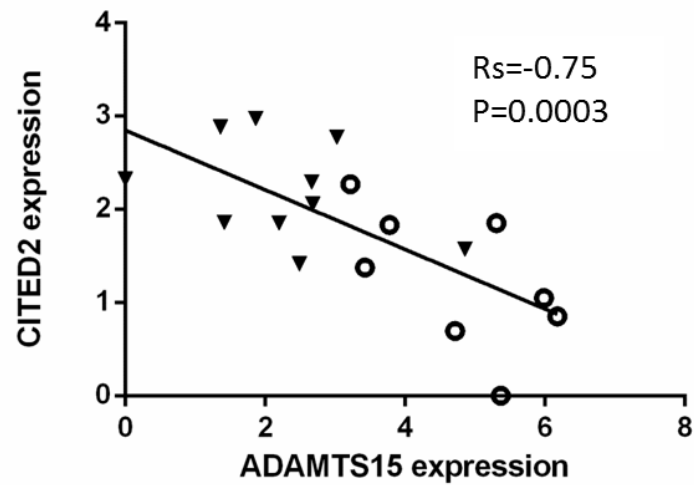
MD-DCs from SpA patients and healthy controls that were left unstimulated (H0), or were stimulated with LPS for 6h or 24h (H6 and H24, respectively) were tested for their capacity to stimulate *in vitro* allogeneic CD4<sup>+</sup> T cells. Results are expressed as <sup>3</sup>H-deoxythymidine incorporation after 6 days of mixed lymphocyte reaction, in CPM ratio. Bars represent the mean proliferation index and standard error of the mean induced by MD-DCs from 19 SpA patients and 24 healthy controls. Wilcoxon tests showed significant differences at H0 (\* p < 0.05) and H6 (\*\* p < 0.01).



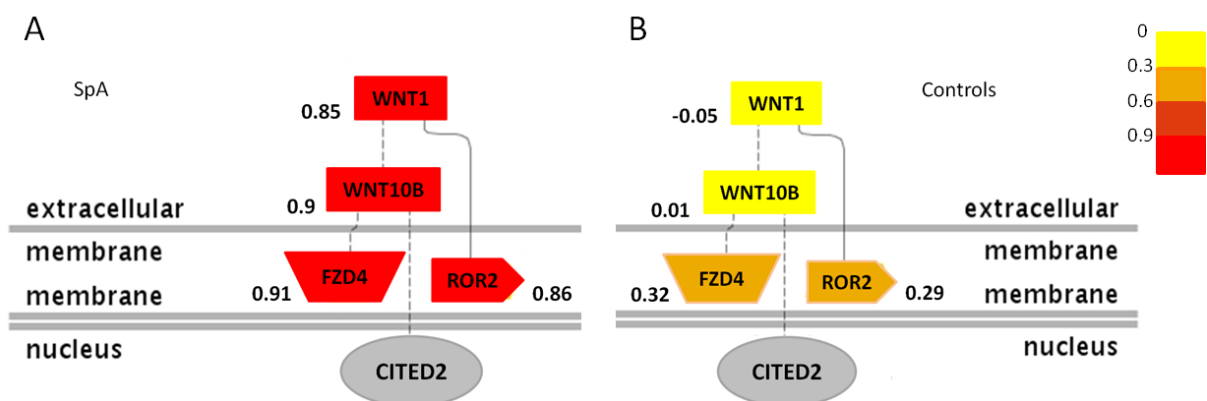
**Figure 2. Unsupervised hierarchical clustering of 57 samples based on microarray expression levels of the 81 differentially expressed genes between SpA patients and controls.** Each time-point is color-coded (H0: blue, H6: red, H24: green). Samples are clustered on the horizontal axis (top row: SpA patients, bottom row: healthy controls), with the length on the vertical axis representing the degree of correlation between samples.



**Figure 3.** Scatter plots showing qRT-PCR expression levels in SpA and controls of the 4 selected genes: *ADAMTS15* (A), *CITED2* (B), *F13A1* (C) and *SELL* (D). The x axis of the plots represents the three LPS-stimulation time-points (H0, H6 and H24) and the y axis shows the log2 of gene expression level normalized with housekeeping genes ( $\Delta\Delta Ct$ ). \* $P < 0.05$ , \*\*\* $P < 0.0005$ , SpA *versus* controls.



**Figure 4. Correlation of qRT-PCR expression levels between *CITED2* and *ADAMTS15*.** Expression is shown for H6 time-point in mixed SpA (open circles) and control (black shapes) samples. P-value (P) and nonparametric Spearman correlation coefficient value (rho) are shown.



**Figure 5. Coexpression of the Wnt signaling pathway with *CITED2* in SpA.** Graphical view of the Wnt pathway identified by studying genes coexpressed at H6 with *CITED2* in the whole microarray dataset in patients (A) and in controls (B). Numbers and color code (scale on the right side) indicate nonparametric Spearman correlation coefficient of the network's gene with *CITED2*

## ADDITIONAL FILES

**Additional file 1: Table S1.** Characteristics of the study patients.\*

Feature	Flow cytometry (n = 10)	Proliferation assay (n = 19)***	Transcriptomic study (n = 9)***
Male gender, n (%)	3 (30)	12 (63)	5 (67)
Age, mean (SD), yrs	43.9 (10.2)	46.4 (9.9)	48.1 (7.9)
HLA-B27, n (%)	10 (100)	18 (94.7)	9 (100)
Disease duration, mean (SD), yrs	23.0 (9.2)	23.2 (10.2)	15.1 (5.2)
BASDAI, mean (SD)	2.9 (1.8)	3.4 (2.2)	3.9 (2.0)
CRP, mean (SD), mg/L	ND	10.6 (14.7)****	7.1 (8.5)
Axial manifestations			
- Inflammatory back pain, n (%)	10 (100)	12 (100)	9 (100)
- Radiographic sacroiliitis**, n (%)	7 (70)	11 (95)	8 (89)
Peripheral manifestations			
- Peripheral arthritis, n (%)	3 (30)	6 (32)	2 (22)
- Enthesitis, n (%)	4 (40)	11 (58)	6 (67)
Extra-articular manifestations			
Uveitis, n (%)	0 (0)	5 (26)	3 (33)
Psoriasis, n (%)	2 (20)	4 (21)	3 (33)
Inflammatory bowel disease, n (%)	0 (0)	2 (10)	0 (0)
Current treatment			
- NSAID, n (%)	3 (30)	14 (74)	5 (56)
- Anti-TNF therapy, n (%)	3 (30)	3 (16)	2 (22)
- None, n (%)	4 (40)	4 (21)	2 (22)

SpA = spondyloarthritis; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; NSAID = non-steroidal anti-inflammatory drug; TNF = tumor necrosis factor. SD = standard deviation; ND = not done.

\* The registered manifestations correspond to those present at the time of examination, or retrieved from past-medical history.

\*\* Refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral.

\*\*\* 6 patients are common to both studies.

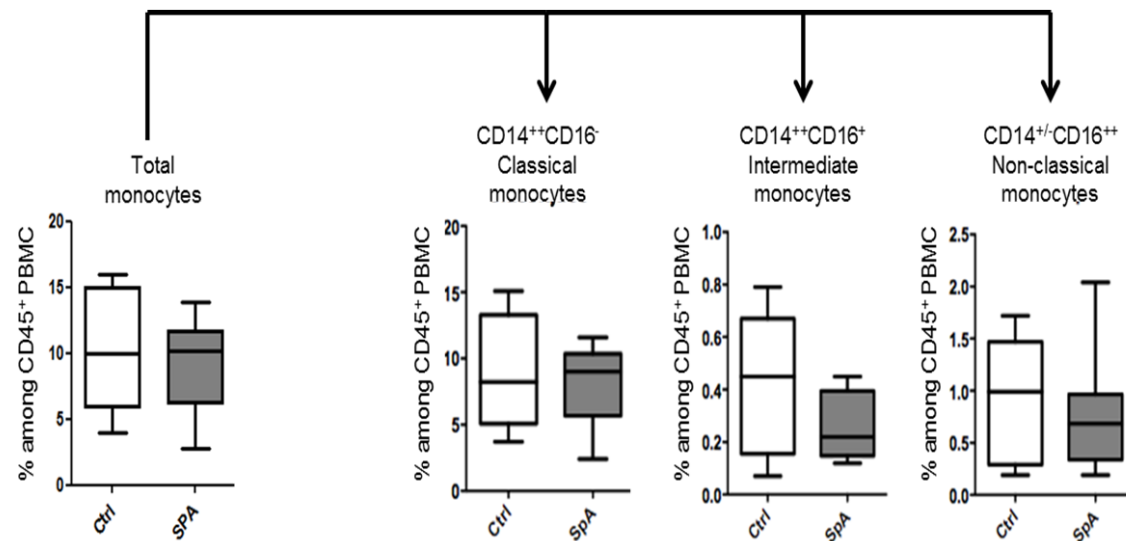
\*\*\*\* Data available for 9 patients.

**Additional file 2: Table S2.** Characteristics of the study healthy controls

Feature	Flow cytometry (n = 5)	Proliferation assay (n = 24)***	Transcriptomic study (n = 10)***
Male gender, n (%)	2 (40)	7 (29.2)	6 (60)
Age, mean (SD), yrs	44 (16.6)	44.5 (13.3)	37.3 (13.3)
HLA-B27, n (%)	0 (0)	0 (0)	0 (0)

**Additional file 3: Table S3.** Nucleotide sequence of the PCR primers

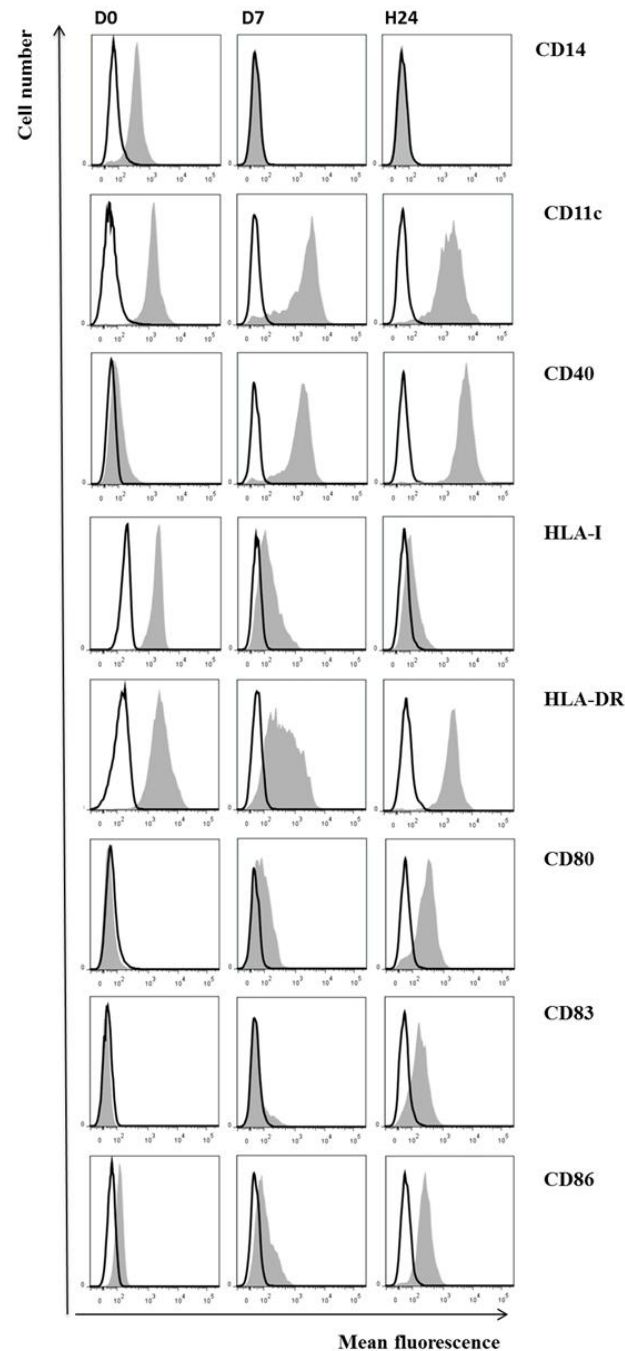
Gene name	Forward primer	Reverse primer
Candidate genes		
<i>ADAMTS15</i>	GACTCTTCACCAAGCCCATGC	AGCCAGGTAGTTGTCATCCCC
<i>CITED2</i>	TGGTGATAGAAATGGGTTTGG	GTTTCGATCGAGTCAACAGC
<i>F13A1</i>	GTGAAGATGATGCTGTGTATCTG	ATGCCATCTTCAAACCTGACC
<i>SELL</i>	CAGTCTACCTGCAGCACAGC	TGGGTGCTCTGACATTTC
Housekeeping genes		
<i>ACTB</i>	GGACTTCGAGCAAGAGATGG	AGGAAGGAAGGCTGGAAGAG
<i>GAPDH</i>	CCTCAACGACCACTTTGTCA	GAGGGTCTCTCTCTTCCTCTTGT
<i>RPL30</i>	CCTAAGGCAGGAAGATGGTG	AATGACCAATTTTCGCTTTGC



**Additional file 4: Figure S4. Comparison of monocyte subsets distribution among PBMCs between SpA patients and healthy controls (Ctrl) by 6-color flow cytometry.**

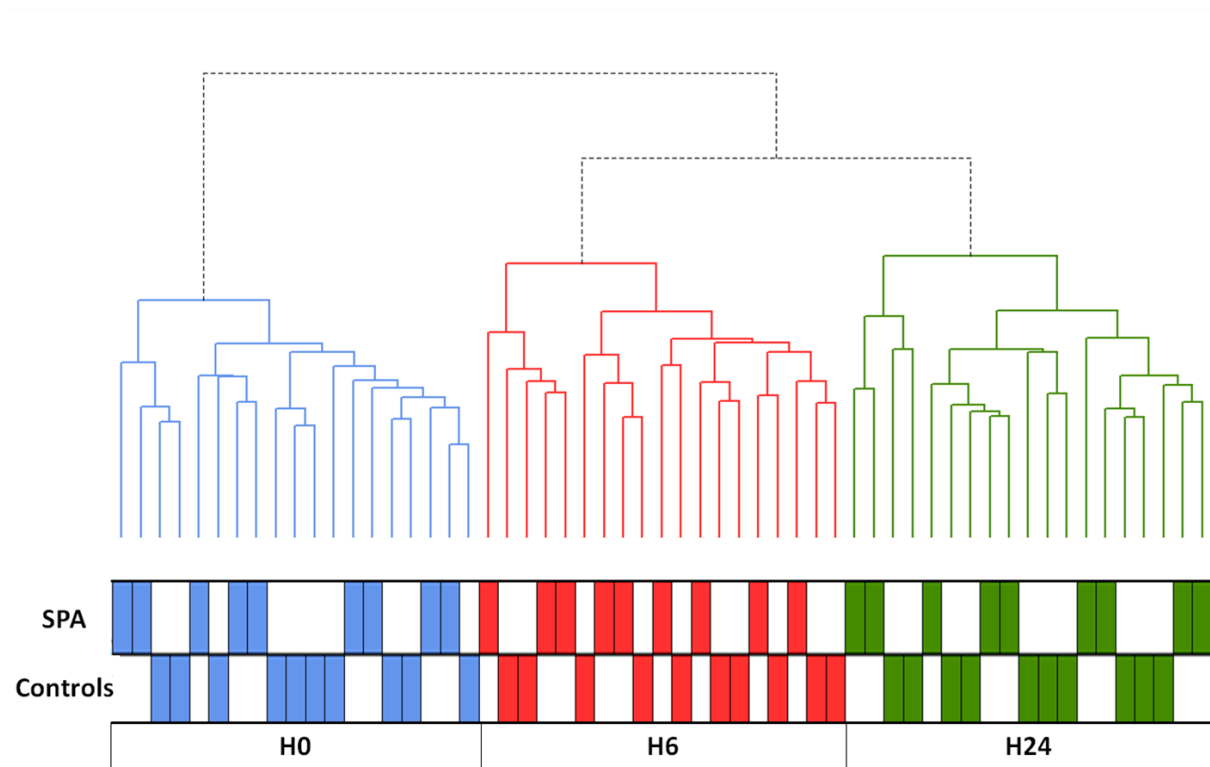
Distribution of monocyte subsets among CD45<sup>+</sup> PBMC was studied in 5 healthy donors (clear boxes) and 10 HLA-B27<sup>+</sup> SpA patients (grey boxes). Results are represented as boxes, bars indicate medians. The mean age of healthy donors was 42 years at the time of the study and 40% of them were men. Characteristics of the patients are shown in Supplementary table 1.





**Additional file 5: Figure S5.** Phenotypic characterization of purified monocytes and MD-DCs by flow cytometry.

Results of one healthy donor representative of study subjects are shown. Monocytes (D0) were CD14<sup>+</sup>, CD11c<sup>+</sup>, CD40<sup>-</sup>, HLA-I<sup>+</sup>, HLA-DR<sup>+</sup>, CD80<sup>-</sup>, CD83<sup>-</sup> and CD86<sup>-</sup>. MD-DCs (D7) were CD14<sup>-</sup>, CD11c<sup>+</sup>, CD40<sup>+</sup>, HLA-I<sup>+</sup>, HLA-DR<sup>+</sup>, CD80<sup>-</sup>, CD83<sup>dim</sup>, and CD86<sup>+</sup>. HLA-DR, CD83 and CD86 expression increased, and CD80 was induced on MD-DCs after LPS stimulation for 24 h (H24). Clear plot representing control isotype and grey plot tested antibody.



**Additional file 6: Figure S6.** Unsupervised hierarchical clustering of 57 samples based on whole-genome gene expression levels.

Each time point is represented by a color (H0: blue, H6: red, H24: green). Samples are clustered on the horizontal axis (top row: SpA patients, bottom row: healthy controls).

**Additional file 7: Figure S7.** List of the genes differentially expressed in MD-DCs between SpA and controls, ranked by fold-change.\*

\* The criteria for the inclusion of genes in this table are described in Patients and Methods.

Global values are limma values. The 4 genes selected for qRT-PCR validation appear bolded.

Gene name	Ensembl ID	P-Value	Global Fold change
<i>Up-regulated genes</i>			
<i>TRBC2</i>	ENSG00000211772	9.67E-03	2.96
<b><i>SELL</i></b>	<b>ENSG00000188404</b>	<b>6.06E-03</b>	<b>2.96</b>
<b><i>F13A1</i></b>	<b>ENSG00000124491</b>	<b>6.12E-05</b>	<b>2.63</b>
<b><i>ADAMTS15</i></b>	<b>ENSG00000166106</b>	<b>3.02E-04</b>	<b>2.38</b>
<i>RND3</i>	ENSG00000115963	8.84E-03	2.36
<i>SIGLEC15</i>	ENSG00000197046	7.98E-04	2.12
<i>HSPA1A</i>	ENSG00000204389	6.03E-03	2.01
<i>TBCK</i>	ENSG00000145348	1.05E-03	1.99
<i>BACE2</i>	ENSG00000182240	8.46E-03	1.85
<i>FBXL4</i>	ENSG00000112234	5.74E-03	1.85
<i>PIAS2</i>	ENSG00000078043	5.35E-03	1.82
<i>KIAA0907</i>	ENSG00000132680	2.26E-03	1.78
<i>SEMA3C</i>	ENSG00000075223	7.04E-03	1.74
<i>HSPH1</i>	ENSG00000120694	5.82E-03	1.74
<i>EDEM3</i>	ENSG00000116406	5.89E-03	1.71
<i>USP40</i>	ENSG00000085982	8.55E-03	1.65
<i>P4HA1</i>	ENSG00000122884	6.44E-03	1.64
<i>DNAJA4</i>	ENSG00000140403	1.92E-03	1.63
<i>FBXO18</i>	ENSG00000134452	2.96E-03	1.54
<i>PTPLA</i>	ENSG00000165996	5.83E-03	1.52
<i>Down-regulated genes</i>			
<i>NDP</i>	ENSG00000124479	3.21E-03	0.13
<i>OLR1</i>	ENSG00000173391	3.36E-03	0.14
<i>FAIM2</i>	ENSG00000135472	6.58E-04	0.14
<i>ZNF804A</i>	ENSG00000170396	8.11E-03	0.23
<i>PLP2</i>	ENSG00000102007	3.42E-03	0.29
<i>CSF3R</i>	ENSG00000119535	5.64E-04	0.32
<i>EIF4H</i>	ENSG00000106682	1.11E-04	0.34
<i>BCKDHA</i>	ENSG00000248098	2.30E-03	0.35
<i>SGMS2</i>	ENSG00000164023	4.13E-05	0.36
<i>GTSF1</i>	ENSG00000170627	6.81E-03	0.36
<b><i>CITED2</i></b>	<b>ENSG00000164442</b>	<b>7.29E-04</b>	<b>0.40</b>
<i>COX7B</i>	ENSG00000131174	5.87E-03	0.40
<i>LRRC4</i>	ENSG00000128594	8.03E-05	0.42
<i>MNDA</i>	ENSG00000163563	3.75E-03	0.44
<i>ANKRD3BP1</i>	ENSG00000214262	8.41E-03	0.44
<i>HAUS1</i>	ENSG00000152240	9.11E-04	0.44

<i>TSPYL5</i>	ENSG00000180543	4.77E-04	0.45
<i>RPL26</i>	ENSG00000161970	3.94E-03	0.46
<i>PIGB</i>	ENSG00000069943	1.58E-03	0.47
<i>RPS15A1</i>	ENSG00000214535	4.65E-03	0.48
<i>CKAP2</i>	ENSG00000136108	6.45E-03	0.50
<i>ALG10B</i>	ENSG00000175548	4.90E-03	0.51
<i>RBBP9</i>	ENSG00000089050	2.01E-03	0.51
<i>PARVG</i>	ENSG00000138964	9.96E-03	0.52
<i>P2RX1</i>	ENSG00000108405	8.45E-06	0.52
<i>RPL10A</i>	ENSG00000198755	7.17E-03	0.53
<i>SPATA20</i>	ENSG00000006282	7.41E-03	0.53
<i>GPR180</i>	ENSG00000152749	9.34E-03	0.54
<i>TTC39C</i>	ENSG00000168234	2.79E-04	0.54
<i>RPS4X</i>	ENSG00000198034	6.85E-03	0.54
<i>ENY2</i>	ENSG00000120533	5.10E-03	0.54
<i>ANAPC15</i>	ENSG00000110200	1.39E-03	0.55
<i>FHL3</i>	ENSG00000183386	7.01E-03	0.55
<i>WDR25</i>	ENSG00000176473	4.45E-05	0.56
<i>FAU</i>	ENSG00000149806	6.03E-04	0.56
<i>TRIM24</i>	ENSG00000122779	1.29E-03	0.57
<i>RFC3</i>	ENSG00000133119	1.03E-03	0.57
<i>TMEM205</i>	ENSG00000105518	4.43E-04	0.57
<i>ITPRIP</i>	ENSG00000148841	7.55E-03	0.58
<i>PORCN</i>	ENSG00000102312	8.47E-03	0.58
<i>BLOC1S1</i>	ENSG00000135441	3.99E-03	0.58
<i>CRTAP</i>	ENSG00000170275	4.34E-03	0.59
<i>ELMO1</i>	ENSG00000155849	6.58E-03	0.59
<i>NUDT3</i>	ENSG00000112664	4.60E-06	0.60
<i>USP30</i>	ENSG00000135093	3.20E-03	0.60
<i>COX20</i>	ENSG00000203667	4.08E-03	0.61
<i>CTBP1-AS1</i>	ENSG00000196810	5.98E-03	0.61
<i>TNFSF13B</i>	ENSG00000102524	3.15E-04	0.61
<i>IFT52</i>	ENSG00000101052	3.79E-03	0.61
<i>RPS15</i>	ENSG00000115268	4.69E-03	0.62
<i>ACADM</i>	ENSG00000117054	9.23E-03	0.62
<i>TFAM</i>	ENSG00000108064	6.40E-04	0.63
<i>MBIP</i>	ENSG00000151332	8.04E-04	0.63
<i>POU5F1B</i>	ENSG00000212993	1.55E-04	0.63
<i>MUT</i>	ENSG00000146085	2.72E-03	0.63
<i>SLU7</i>	ENSG00000164609	1.50E-04	0.63
<i>BAK1P1</i>	ENSG00000175730	4.24E-03	0.64
<i>ZFP36L2</i>	ENSG00000152518	9.55E-03	0.65
<i>SAP130</i>	ENSG00000136715	1.15E-04	0.65
<i>POLRID</i>	ENSG00000186184	2.24E-04	0.65
<i>FAM204A</i>	ENSG00000165669	7.35E-05	0.66

---

## Article 5

---

### ***ERAP1* Gene Expression is influenced by Non-Synonymous Polymorphisms associated with Predisposition to Spondyloarthritis**

*Félicie Costantino\**, Alice Talpin\*, Irini Evnouchidou, Amir Kadi, Ariane Leboime, Roula Said-Nahal, Nelly Bonilla, Franck Letourneur, Tifenn Leturcq, Zeyna Ka, Peter van Endert, Henri-Jean Garchon, Gilles Chiochia, Maxime Breban

*\*Les deux premiers auteurs ont contribué de manière équivalente à ce travail.*

Article soumis à « Arthritis & Rheumatology » en septembre 2014.

#### **L'expression du gène *ERAP1* est influencée par les polymorphismes non synonymes associés à la prédisposition à la spondyloarthrite**

Ce travail s'est inscrit dans la suite de l'étude du transcriptome des MD-DCs décrite dans l'Article 4<sup>415</sup>. En effet, le gène *ERAP1* figurait parmi les gènes différenciellement exprimés entre patients et témoins. Sans atteindre le seuil de significativité statistiques fixé dans cette étude ( $p < 0,01$ ), on observait une augmentation de son taux d'expression chez les patients par rapport aux témoins principalement après 6 heures de stimulation par le LPS ( $p = 0,03$ ).

Le gène *ERAP1* code pour une aminopeptidase du réticulum endoplasmique dont la fonction principale d'apprêtement des peptides avant leur présentation par les molécules du CMH de classe I peut aisément être reliée à la SpA. Plusieurs polymorphismes d'*ERAP1* ont été associés au risque de SA puis à celui de SpA dans son ensemble. Toutefois, compte-tenu du DL génétique important qui existe dans la région entourant *ERAP1*, le polymorphisme causal n'a pu être formellement identifié. Le SNP rs30187 semble être le principal responsable du signal d'association avec un signal secondaire porté par les SNPs rs10050860 et rs17482078<sup>269</sup>. La combinaison de ces trois SNPs (rs17482078/rs10050860/rs30187) détermine la formation de trois haplotypes majoritaires

conférant différents niveaux de susceptibilité à la SpA : un haplotype à risque (CCT), un haplotype neutre (CCC) et un haplotype protecteur (TTC)<sup>383</sup>.

Les études fonctionnelles portant sur ces polymorphismes ont montré que l'allèle de susceptibilité de rs30187 était associée à une augmentation de l'activité enzymatique d'ERAP1, tout comme l'haplotype à risque CCT<sup>269,357,359,360</sup>. Au vu des résultats d'expression obtenus lors de notre étude du transcriptome des MD-DCs, nous avons recherché s'il existait une association entre les polymorphismes d'*ERAP1* associés à la SpA et le niveau d'expression du gène.

Nous avons tout d'abord montré que dans la cohorte étudiée dans l'**Article 4** (constituée de neuf patients ; dix témoins), l'expression d'ERAP1 dans les MD-DCs laissées au repos ou stimulées par du LPS était fortement associée au génotype de rs30187, avec une expression augmentée chez les porteurs de l'allèle de susceptibilité T. Un score a été attribué à chaque sujet testé en fonction du niveau de susceptibilité à la SpA des haplotypes qu'il possédait. Il existait une forte association entre ce score de risque haplotypique et le niveau d'expression d'ERAP1. Cette association a été confirmée dans une deuxième cohorte incluant 14 patients et 34 témoins (13 individus HLA-B27+ appartenant à la fratrie des malades et 21 témoins indépendants dont dix étaient HLA-B27+). L'ensemble des associations mises en évidence était indépendante de la maladie ou de la présence du HLA-B27. Nous nous sommes ensuite intéressés aux individus apparentés de la cohorte de réplique (13 couples de germains, tous HLA-B27+, l'un malade, l'autre non malade). Au sein des couples partageant le même haplotype d'*ERAP1*, on observait une forte corrélation de l'expression d'ERAP1, alors qu'il n'existait aucune corrélation dans les fratries haplo-discordantes.

Afin de chercher à évaluer les conséquences d'une augmentation de l'expression de l'ARN messager d'ERAP1, nous avons ensuite étudié l'expression protéique d'ERAP1 en fonction des haplotypes portés par les individus. Pour ce faire, des lignées lymphoblastoïdes B provenant de patients homozygotes pour chacun des haplotypes d'*ERAP1* précédemment décrits ont été obtenues par immortalisation par le virus EBV. L'expression protéique d'ERAP1 dans ces lignées, évaluées par Western blot, était corrélée au score de risque haplotypique et au niveau d'expression du transcrit. En ce qui concerne l'activité enzymatique, on notait une différence entre la lignée porteuse de l'haplotype protecteur et les autres lignées, sans que cette tendance ne soit statistiquement significative. Une tendance similaire a été observée dans les MD-DCs provenant d'un sous-groupe des patients précédemment inclus dans l'étude transcriptomique et qui ont été prélevés à nouveau pour cette raison.

L'ensemble de ces résultats permet de conclure que les polymorphismes d'*ERAP1* associés à la SpA influencent le niveau d'expression du gène. La modulation du niveau d'expression génique

d'ERAP1 semble également avoir des conséquences sur son expression protéique et sur son niveau d'activité enzymatique. Il reste néanmoins à déterminer les conséquences d'une augmentation de l'activité enzymatique d'ERAP1 sur la susceptibilité à la SpA.

# **ERAP1 Gene Expression is influenced by Non-Synonymous Polymorphisms associated with Predisposition to Spondyloarthritis**

Félicie Costantino, MD,<sup>1,2\*</sup> Alice Talpin, PhD,<sup>1\*</sup> Irini Evnouchidou PhD,<sup>3</sup> Amir Kadi, PhD,<sup>1</sup>, Ariane Leboime, MD,<sup>2</sup> Roula Said-Nahal, MD,<sup>2</sup> Nelly Bonilla, MS,<sup>1</sup> Franck Letourneur, MS,<sup>4</sup> Tifenn Leturcq, MD,<sup>1</sup> Zeyna Ka, MS,<sup>1</sup> Peter van Endert, MD, PhD,<sup>3</sup> Henri-Jean Garchon, MD, PhD,<sup>1,5</sup> Gilles Chiochia, PhD,<sup>1</sup> Maxime Breban, MD, PhD<sup>1,2</sup>

*1 – INSERM U987, UFR Simone Veil, Versailles-Saint Quentin University, Laboratoire d'Excellence INFLAMEX, France*

*2 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

*3 – INSERM U1151 ; CNRS UMR 8253 ; Université Paris Descartes, Sorbonne Paris Cité, Paris, France*

*4 – Institut Cochin, INSERM U1016, CNRS (UMR8104), Paris Descartes University, Paris, France*

*5 – Genetics Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, France*

*\* The two first authors contributed equally to that work.*

This work was supported by a grant from Agence Nationale de la Recherche (grant ANR 2010 GEMISA). Félicie Costantino was supported by a grant from the Société Française de Rhumatologie (SFR). Alice Talpin was supported by a grant from Arthritis Fondation Courtin. Irini Evnouchidou was supported by a Thyssen Foundation grant.

We acknowledge the financial support of the Investissements d'Avenir programme ANR-11-IDEX-0005-02, Sorbonne Paris Cité, Laboratoire d'excellence INFLAMEX, and special support for equipments of the Communauté d'Agglomération CASQY, and Conseil Général des Yvelines

Address reprint requests and correspondence to:

Maxime Breban, M.D., Ph.D., Rheumatology Division

Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France

Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865

E-mail: [maxime.breban@apr.aphp.fr](mailto:maxime.breban@apr.aphp.fr)

**Keywords:** ankylosing spondylitis; spondyloarthritis; single-nucleotide polymorphism; genetics; ERAP1; eQTL.

Word Count: 3,443



## ABSTRACT

**Objective:** Several polymorphisms in *ERAP1* are strongly associated with susceptibility to spondyloarthritis (SpA). The combination of rs17482078, rs10050860 and rs30187 determines three major haplotypes with different levels of association with SpA (“protective” T/T/C; “neutral” C/C/C and “susceptibility” C/C/T). The aim of the present work was to determine whether such haplotypes might affect *ERAP1* mRNA expression, protein level and/or enzymatic activity in antigen-presenting cells, a type of cell potentially relevant to disease pathogenesis.

**Methods:** Monocyte-derived dendritic cells (MD-DCs) were generated in two cohorts comprising a total of 23 SpA patients and 44 healthy controls. B lymphoblastoid cell lines (B-LCLs) were established from individuals homozygous for risk, neutral or protective *ERAP1* haplotype respectively. In those samples, we investigated the relation between *ERAP1* haplotypes and mRNA expression level. We also measured the relative protein expression of ERAP1 by Western blot and its enzymatic activity using a fluorogenic assay.

**Results:** In MD-DCs, there was a strong association between *ERAP1* haplotypes and *ERAP1* mRNA expression level, with higher levels in subjects harboring the susceptibility haplotype in both cohorts ( $p = 0.001$  and  $5.6 \times 10^{-7}$ , respectively). In B-LCLs, we observed a significant correlation between haplotype risk score and *ERAP1* transcript or protein level ( $p = 0.003$ ,  $\rho = 0.92$  for both). Enzymatic activity followed a similar trend both in MD-DCs and in B-LCLs.

**Conclusions:** Those data provide strong evidence that SpA-associated *ERAP1* polymorphisms affect the level of gene expression in antigen-presenting cells. How an increased production/activity of ERAP1 may influence susceptibility to SpA remains to be determined.

## INTRODUCTION

Spondyloarthritis (SpA) is a frequent form of chronic inflammatory arthritis with an estimated life-time risk around 0.43% in adult Caucasian population (1). It comprises several closely related disorders which share clinical and imaging features as well as genetic predisposition (2,3). Familial aggregation and heritability of SpA are high and involve multiple genetic factors among which HLA-B27 allele plays a prominent role (4).

The first genome-wide association study performed in ankylosing spondylitis (AS), the prototypical form of SpA, revealed significant associations of several non-synonymous single-nucleotide polymorphisms (SNPs) situated in the endoplasmic reticulum (ER) aminopeptidase 1 gene (*ERAP1*) with the disease (5). Such associations were subsequently confirmed in independent Caucasian and Asian cohorts of AS (6-8) as well as with the whole group of SpA, beyond AS (9,10). Interestingly, the association between *ERAP1* polymorphisms and SpA seems to be restricted to HLA-B27 positive disease (10-12). More recently, new associations have been described between AS and three other aminopeptidases (*ERAP2*, *LNPEPP* and *NPEPPS*) (12).

The *ERAP1* gene located on chromosome 5q15 encodes an aminopeptidase expressed in the ER. Its main documented function is to trim peptides to the optimal length for binding to major histocompatibility class I (MHC-I) molecules (13,14). It has also been involved in the shedding of cell surface pro-inflammatory cytokine receptors (TNF-R1, IL6-R, IL1-R2) (15-17). However such sheddase activity remains debated, since *erap1* knock-out mice exhibit levels of soluble TNF-R and IL6-R similar to wild-type animals (11). Moreover AS-associated *ERAP1* polymorphisms do not seem to influence the serum levels of soluble cytokine receptors among patients with AS (18). These results together with a restriction of the association of *ERAP1* to HLA-B27-positive disease suggest that the pathological effects of *ERAP1* in AS could be related to its trimming peptides function in the ER with potential consequences on antigen presentation by MHC-I molecules.

Due to an extensive linkage disequilibrium (LD) existing within and surrounding the *ERAP1* gene locus, it appears difficult to identify with certainty causal polymorphism(s) (19). A two-mutations model has been proposed by *Evans et al.* with a primary effect due to rs30187 and a secondary effect due to either rs10050860 or rs17482078 (11). This model is fully consistent with studies showing a strongly significant association of a haplotype combining those 3 polymorphisms, with AS in Canada and in Portugal and with SpA in

France and Belgium (6,9,20). However functional consequences of bearing this haplotype and their relevance for SpA susceptibility remain to be elucidated.

Several studies using recombinant ERAP1 mutants and synthetic peptides showed a decreased enzymatic activity of ERAP1 in the presence of AS-protective alleles of rs30187 and rs17482078 (11,21-25). It has been suggested that the reduced activity observed with rs30187 protective allele (K528R) was linked to modifications of ERAP1 structure (21,26). Another hypothesis is that *ERAP1* polymorphisms could affect gene expression level (19). In the present study, we investigated an influence of the ERAP1 genetic makeup on its quantitative expression, at both messenger and protein level, also including enzymatic activity, in the monocyte-derived dendritic cell (MD-DC), an antigen-presenting cell relevant to study ERAP1 function and potentially involved in SpA pathogenesis (27-32). We also accounted for the marked LD. We uncovered a tight relationship between ERAP1 quantitative expression and the ERAP1 locus-associated genetic risk.

## **PATIENTS & METHODS**

### **Patients and controls recruitment**

All participants of the study gave written informed consent and the study has been approved by the local ethical committee of Ile de France XI, Poissy-Saint-Germain Hospital (Saint-Germain-en-Laye, France). All the patients satisfied the Assessment of SpondyloArthritis international Society classification for SpA (33,34). Detailed clinical characteristics of the enrolled patients are provided in **Table 1**.

The control populations consisted of sex- and age-matched independent healthy donors (“discovery” and “replication” cohorts) and unaffected siblings of patients (“replication” cohort).

### **Cells isolation and culture**

#### *Peripheral blood monocytes isolation*

Peripheral blood mononuclear cells were isolated from 50 mL of freshly drawn blood by gradient separation on Ficoll density gradient centrifugation (Stemcell Technologies, Vancouver, Canada). Monocytes were obtained by magnetic cell sorting using anti-CD14 monoclonal antibody (mAb) coated beads (BD Biosciences, San Jose, CA, USA). Sorted

monocytes were morphologically homogeneous with 99% of CD14<sup>+</sup> cells, as determined by flow cytometry (32).

#### *MD-DCs generation and stimulation*

To obtain MD-DCs, monocytes were further cultured for 6 days in RPMI 1640 medium supplemented with 10% heat inactivated fetal bovine serum, 100 U/mL penicillin, 100 µg/mL streptomycin, 500 U/mL recombinant human granulocyte-macrophage colony stimulating factor (rhGM-CSF) and 500 U/mL rhIL-4 (AbCys S.A., France).

Then, the MD-DCs were stimulated or not with lipopolysaccharide (LPS) from *E. Coli* (Sigma-Aldrich, St. Louis, MO, USA) at a concentration of 100 ng/mL for the last 6 or 24 hours of culture (further referred to as time-points H0, H6 and H24).

#### *Lymphoblastoid B cell lines (B-LCLs) generation*

B-LCLs were established from EBV-transformed lymphocytes according to standard protocol, and grown at 37°C in RPMI 1640 medium supplemented with 10% heat inactivated fetal bovine serum, 100 U/mL penicillin and 100 µg/mL streptomycin.

#### **Genotyping of rs30187, rs17482078 and rs10050860**

Genotypes of rs30187, rs17482078 and rs10050860 were determined through sequencing. Briefly, three amplicons (one for each SNP) were amplified from 200 ng of genomic DNA. Approximately 10 to 30 ng of the purified products were sequenced on ABI3730xl capillary sequencer (Applied Biosystems, Foster City, CA, USA) and the resulting sequences were analyzed using SeqScape software version 2.5 (Applied Biosystems).

#### **LD assessment**

Genotype data were obtained through Immunochip custom array (Illumina, San Diego, California, USA), as part of the International Genetics of AS Consortium (IGAS) study (12). We used a subset of 126 SNPs selected on the following criteria: located from 95.999 to 96.387 Mb (GRCh37) of the chromosome 5p-telomere (spanning *CAST*, *ERAP1*, *ERAP2* and *LNPEP* genes region); minor allelic frequency > 0.2 and pairwise  $r^2$  > 0.8. LD for reconstructed haplotypes was determined using the Haploview program V4.2 (35).

### ***ERAP1* mRNA expression**

Total RNA was isolated with RNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. *ERAP1* mRNA expression level was then evaluated with micro-array hybridization. Quantitative real-time polymerase chain reaction (qRT-PCR) was also performed in the “discovery” cohort of MD-DCs samples for technical validation.

#### *Gene expression micro-array*

For micro-array hybridization, RNA was reverse transcribed, converted to biotinylated complementary RNA using standard Affymetrix protocol and hybridized to the Affymetrix GeneChip Human Gene 1.0 ST Array (Affymetrix, Santa Clara, CA, USA). Raw data were then transformed by Robust Multichip Analysis (RMA) (36).

#### *qRT-PCR*

For qRT-PCR, RNA was reverse-transcribed using Superscript II (Invitrogen, Carlsbad, CA, USA). A 147 bp region common to all the known protein-coding *ERAP1* transcripts was amplified with the following primers (Forward (F): TTTGAACTTGGCTCATCTTCC; Reverse (R): AATTGTCTGTTGGACACAACG) and then quantified using the SYBR green PCR Master Mix (Applied Biosystems) and the 7300 Real-Time PCR System (Applied Biosystems). Gene expression data were computed with the  $\Delta\Delta C_q$  method, using three housekeeping genes: *RPL30* (F: CCTAAGGCAGGAAGATGGTG; R: AATGACCAATTTGCTTTGC),  $\beta$ -*ACT* (F: GGACTTCGAGCAAGAGATGG; R: AGGAAGGAAGGCTGGAAGAG) and *GAPDH* (F: CCTCAACGACCACTTTGTCA; R: GAGGGTCTCTCTTCTCCTCTTGT).

### **ERAP1 protein expression (Western-blot)**

Cells were lysed on ice in lysis buffer. Cell lysates were centrifuged at 12,000 rpm for 15 min. 28  $\mu$ g of protein, as determined by Pierce BCA protein assay kit (Thermoscientific, Waltham, MA, USA), was suspended in reducing Laemmli Buffer 2X and heated at 95°C for 5 min. Proteins were separated on a 7.5% SDS polyacrylamide gel (Biorad, Hercules, CA, USA) and transferred to polyvinylidene difluoride membrane. Blocking was for 1 h at room temperature (RT) with PBS 1X, 0.01% Tween 20, containing 5% dry milk, followed by incubation with an anti-ERAP1 mAb (Abcam, Cambridge, MA, USA) diluted at 1:2,000 for 2 h at RT and with HRP-conjugated secondary antibody (goat anti-rabbit IgG, Dako, Ft. Collins, CO, USA) diluted at 1:10,000 for 1 h at RT. Immunostaining of the blot with an anti-beta-actin mAb

(Santa Cruz Biotechnology, Dallas, TX, USA) served as loading control. Labeled proteins were detected using enhanced chemiluminescence reagent (Amersham Biosciences, UK) and the signal read with the Fusion FX7 system (Vilber Lourmat, Marne-la-Vallée, France). Band intensities were quantified by densitometry analysis using ImageJ software (NIH, Bethesda, MD, USA).

### **ERAP1 enzymatic assay**

#### *ERAP1 immunoprecipitation*

Cell pellets ( $25 \times 10^6$  cells for B-LCLs or  $2 \times 10^6$  cells for MD-DCs) were lysed in 125  $\mu$ l or 50  $\mu$ l of lysis buffer, respectively and left on ice for 30 min. The lysate (supernatant) was recovered by centrifugation at 14,000 rpm for 10 min at 4°C. Before immunoprecipitation, the total protein content was quantified via Bradford assay. The lysate was then precleared by incubation with Glycine sepharose beads for 1 h before incubation for at least 2 h with ERAP-specific mAb 4D2-coupled conjugated to sepharose beads. The beads were then washed 3 times in buffer (50 mM Tris, 150 mM NaCl, pH 7.4) and resuspended in 500  $\mu$ l buffer before being used in activity assay.

#### *ERAP1 enzymatic activity measurement*

The fluorogenic substrate L-AMC (50  $\mu$ l of 100  $\mu$ M in 50mM Tris pH 7.4, 150 mM NaCl) was incubated with 50  $\mu$ l of immunoprecipitated ERAP1 at 25 °C and the fluorescence signal was measured for 15 min on a Mithras LB 940 plate reader (Berthold Technologies, Thoiry, France) with excitation set at 355 nm and emission at 460 nm. The hydrolysis rate of the substrate was calculated via the resulting time slope. The error bars represent the standard deviation of 3 measurements.

### **MHC- I and HLA-B27 surface expression**

For assessment of MHC-I and HLA-B27 cell surface expression, MD-DCs were incubated for 20 min with FITC-labelled anti-HLA-I or anti-HLA-B27 IgG mAbs, (BD Biosciences), at 4°C. Then, cells were washed with PBS 1X and fixed with 2% formaldehyde. The analyses were performed by flow cytometry on a FACS canto II apparatus (BD Biosciences).

### **Statistical analysis**

Association between haplotype score and SpA susceptibility was tested using logistic regression analysis. Multivariate analysis of variance (MANOVA) was used to test an

association between *ERAP1* gene expression in MD-DCs samples and *ERAP1* genotypes/haplotypes, disease status and HLA-B27 positivity at all stimulation time-points. In case of statistical significance ( $p < 0.05$ ), post-hoc one-way analysis of variance (ANOVA) was performed at each stimulation time-point. Correlation between *ERAP1* expression levels in sibling-pairs was tested using a linear model with mixed effects followed by a likelihood ratio test comparing the full model and the null model to assess the significance of the correlation. A non parametric Spearman test was used to test correlation between haplotype risk score and *ERAP1* transcript or protein levels in B-LCLs.

## RESULTS

### Association between rs30187 genotype and *ERAP1* gene expression

The primary association signal of *ERAP1* with AS susceptibility is assumed to be borne by rs30187 (11). Thus, we first studied the influence of this SNP on *ERAP1* gene expression in two independent case-control cohorts. The discovery cohort included 9 HLA-B27 positive patients with axial SpA (of whom 8 had AS) and 10 unrelated healthy donor controls (32). The replication cohort consisted of 14 HLA-B27<sup>+</sup> SpA patients (including 9 with AS) and 34 healthy controls (13 HLA-B27<sup>+</sup> siblings of patients and 21 additional independent controls (10 HLA-B27<sup>+</sup> and 11 HLA-B27<sup>-</sup>)).

In the discovery cohort, rs30187 genotypes were significantly associated with *ERAP1* expression in MD-DCs, at each stimulation time-point, as determined by micro-arrays with an additive effect (combined  $p = 9.0 \times 10^{-4}$ ). Indeed, expression levels increased with the number of AS-susceptibility T alleles (**Figure 1A**). This result was confirmed by qRT-PCR on the same samples (combined  $p = 2.0 \times 10^{-4}$ ; **Figure 1C**) and further replicated in the replication cohort (**Figure 1B**) in which it was observed in both circulating monocytes ( $p = 0.001$ ) and MD-DCs (combined  $p = 7.4 \times 10^{-10}$ ). Such association was independent of disease or HLA-B27 status (data not shown).

No such association between rs30187 genotypes and either *ERAP2* or *LNPEP* expression levels was observed in the discovery cohort ( $p = 0.54$  and  $p = 0.24$ , respectively). Moreover, we did not find any association between the two other SpA-associated *ERAP1* polymorphisms studied, i.e. rs17482078 nor rs10050860 and *ERAP1* transcript levels ( $p = 0.28$  for both SNPs which were in perfect LD with each other).

### Extended LD around the *ERAP1* locus

The major protective non-ancestral allele of rs30187 results in the substitution of a lysine for an arginine (K528R). As regulatory variants are rarely located in coding regions, rs30187 may be in LD with one or several other variants located in regulatory elements. To assess the extent of the LD, we used genotype data from a French collection of 354 AS patients and 346 healthy controls included in the IGAS Immunochip study (12). We extracted a set of 126 tag-SNPs covering the region surrounding *ERAP1*, from *CAST* to *LNPEP* genes. A high LD was observed throughout this region with two major blocks, one including *ERAP1* and the other including *ERAP2* and *LNPEP* (**Figure 2A**). Consequently, numerous variants could be responsible for *ERAP1* gene expression variation.

*ERAP1* genetic studies have suggested that there is a secondary association signal with AS, due either to rs10050860 or to rs17482078. Combination of these 3 SNPs resulted in 3 major rs17482078/rs10050860/rs30187 haplotypes: T/T/C (20.1%), C/C/C (40.1%) and C/C/T (39.0%) (**Figure 2B**). Thus studying rs17482078/rs10050860/rs30187 haplotypes rather than rs30187 alone could improve the detection of gene expression variation depending on genetic sequence.

### Association between rs17482078/rs10050860/rs30187 haplogenotypes and *ERAP1* gene expression

We have previously reported that rs17482078/rs10050860/rs30187 haplotypes were not equally distributed between SpA patients and controls: the C/C/T haplotype was more frequent in patients, T/T/C was overrepresented in controls, whereas C/C/C was evenly distributed (9). Accordingly, we attributed a score to each haplotype corresponding to its level of association with SpA (“protective” T/T/C = -1; “neutral” C/C/C = 0 and “susceptibility” C/C/T = 1). We thus determined a score for each of the 700 French subjects included in the IGAS Immunochip study (12), corresponding to the sum of its 2 haplotypes scores. There was indeed a significant association of such score with SpA susceptibility, allowing us to consider it as “risk” score (**Supplementary table 1**;  $p = 0.002$ ; odds ratio = 1.27; 95% confidence interval: 1.1 to 1.48).

In both cohorts, we observed a highly significant positive correlation between the haplotype risk score and *ERAP1* expression level (**Figure 3A,B**). This result was consistently observed in MD-DCs either stimulated or not (combined  $p = 0.001$  and  $5.6 \times 10^{-7}$  in the



discovery and replication cohorts, respectively) and in monocytes ( $p = 9.9 \times 10^{-6}$ ; **Figure 3B**) and without any detectable influence of disease or HLA-B27 status.

Transcript abundances are known to be largely heritable (37). In the 13 sibling-pairs included in the replication cohort, we observed a significant correlation of *ERAP1* expression between siblings ( $p = 1.2 \times 10^{-4}$ ). Interestingly, in the 6 sibling-pairs haploidentical for rs17482078/rs10050860/rs30187 haplotypes, such correlation was dramatically high ( $p = 4 \times 10^{-7}$ ;  $r^2 = 0.84$ ; **Figure 4A**) whereas it was of borderline significance in the 7 sibling-pairs carrying a divergent haplotypes ( $p = 0.05$ ;  $r^2 = 0.11$ ; **Figure 4B**).

### **Association between *ERAP1* rs17482078/rs10050860/rs30187 haplotype, mRNA level, protein level and enzymatic activity in EBV-cell lines**

To further investigate the functional role of *ERAP1* haplotypes at the transcription and protein levels, we studied 7 HLA-B27+ patients homozygous for either one of each haplotype: 3 carrying the C/C/T susceptibility haplotype, 3 carrying the C/C/C neutral one and 2 carrying the T/T/C protective one. B-LCLs from those patients were immortalized with EBV. One additional HLA-B27+ healthy control homozygous for the T/T/C haplotype was included in the enzymatic activity assay.

In those cell lines, we observed a significant correlation between rs17482078/rs10050860/rs30187 haplotype risk score and *ERAP1* transcript levels ( $p = 0.003$ ,  $\rho = 0.92$ ; **Figure 5A**), as in the foregoing studies on monocytes and MD-DCs. A similar correlation was also observed between haplotype risk score and *ERAP1* protein level ( $p = 0.003$ ;  $\rho = 0.92$ ; **Figure 5C**). Although, there was no significant correlation between haplotype risk score and *ERAP1* enzymatic activity, the low risk haplotype cell lines had a lower activity than other B-LCLs (**Figure 5E**). We observed a similar trend in unstimulated MD-DCs samples generated from 8 patients belonging to the discovery and replication cohorts, which were bled again for this purpose (**Figure 5B,D,F**).

### **Consequences of *ERAP1* expression on MHC-I surface expression.**

To investigate putative consequences of *ERAP1* expression changes on antigen-presentation, we measured MHC-I and HLA-B27 expression at the MD-DCs surface. No correlation was observed between *ERAP1* expression level and either whole MHC-I or HLA-B27 expression, whatever the stimulation time-point (data not shown).

## DISCUSSION

Genetic association studies have highlighted the importance of *ERAPI* in SpA susceptibility. However, few mechanisms have been hypothesized to explain the consequences of SpA-associated polymorphisms of this gene. We provide here for the first time strong evidence that *ERAPI* expression was increased at the mRNA level in subjects harboring SpA-susceptibility polymorphisms.

Except for few recent studies (24,25,38), all others have investigated the functional consequences of SpA-associated SNPs of *ERAPI* one by one (11,21-23,39). Using a large case-control cohort from the Immunochip study, we identify a strong LD in the *ERAPI* region as already described (19). Moreover *Evans et al.* have suggested that the pattern of association with AS was not fully captured by studying a single SNP (11), as confirmed by other studies demonstrating association between *ERAPI* haplotypes and SpA (6,7,9). In particular, *Kadi et al.* have identified three distinct rs17482078/rs10050860/rs30187 haplotypes that were differentially associated with SpA: the -TTC haplotype was associated with a low disease risk, -CCC with a neutral risk and -CCT with an increased risk (9). Thus, we decided to study haplotype effects to better understand polymorphisms consequences on disease susceptibility.

Our results obtained on peripheral blood monocytes and MD-DCs demonstrate the presence of a cis-expression quantitative trait locus controlling *ERAPI* expression associated with the rs17482078/rs10050860/rs30187 haplotype. Such control has already been suggested for several *ERAPI* SNPs in cell lines (40) and in skin biopsies (41) but never in primary monocytes or MD-DCs. Recently, *Seregin et al.* obtained different results with no difference in *ERAPI* gene expression between the -TTC and -CCT rs17482078/rs10050860/rs30187 haplotypes (25). However those results were obtained in transfected cell lines overexpressing *ERAPI* and thus may not reflect physiological conditions. Our results were also quite different from those of *Fairfax and al.* in primary monocytes (42). Indeed, they identified rs30187 as a SNP associated with *ERAP2* expression whereas in our study we observed no such association.

Furthermore, we found that *ERAPI* haplotypes affected protein abundance and enzymatic activity. Thus *ERAPI* can also be considered as a protein quantitative trait locus under the cis-regulatory control of the rs17482078/rs10050860/rs30187 haplotype. An increased expression of *ERAPI* protein has been previously reported in AS patients, as compared to controls (43). Such observation could presumably be related to imbalanced

distribution of *ERAP1* haplotypes between cases and controls in that study, as we did not find ourselves any correlation between disease status and *ERAP1* expression, independently of *ERAP1* genotypes.

Interestingly, *ERAP1* gene expression seemed to be well correlated to SpA predisposition. Indeed, expression was low in low-risk haplotype carriers and high in high-risk ones. Several hypotheses can be made to explain why *ERAP1* overexpression might increase SpA susceptibility. First, our results suggested that increased expression of *ERAP1* transcript may lead to an increased enzymatic activity. This increased enzymatic activity had potential consequences on HLA-B27 peptidome, as discussed below. Secondly, *Goto et al.* have highlighted another function of ERAP1 outside the ER. Indeed, they showed in a murine model that ERAP1 was secreted from macrophages in response to activation by treatment with LPS and interferon- $\gamma$ , resulting in heightened phagocytic activity of macrophages (44).

*ERAP1* genetic association with AS is restricted to HLA-B27 positive patients, suggesting a functional interaction between SpA-associated polymorphisms of *ERAP1* and HLA-B27. At variance with other studies we did not find any correlation, either positive or negative, between HLA-I nor HLA-B27 cell surface expression and *ERAP1* expression levels (45). However, SpA-associated SNPs may bear “qualitative” consequences rather than quantitative ones and may affect HLA-B27 peptidome. Consistent with such hypothesis, several studies demonstrated that *ERAP1* polymorphisms could affect HLA-B27 peptidome through their influence on ERAP1 enzymatic activity (25,38,46,47). Higher enzymatic activity leads to an increased destruction of some HLA-B27-restricted epitopes whereas other epitopes are preferentially generated (38,47). Peptidome modifications may in turn bear consequences relevant for SpA susceptibility, such as skewing an HLA-B27-restricted immune response towards a particular antigen (48) or influencing HLA-B27 folding and stability with potential consequences on misfolding or on cell surface expression of heavy chains homodimers (49).

In conclusion, our results demonstrate a strong correlation between SpA susceptibility conferred by the rs17482078/rs10050860/rs30187 haplotypes and *ERAP1* expression levels. Increased gene expression observed in high-risk haplotype carriers was correlated to increased protein expression. Enzymatic activity seemed also to be at least partly influenced by gene expression level together with other putative mechanisms such as structural modifications.

Nevertheless, the consequences of ERAP1 enzymatic activity variations on disease susceptibility remain to be fully elucidated.

## REFERENCES

1. Costantino F, Talpin A, Said-Nahal R, Goldberg M, Henny J, Chiocchia G, et al. Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann Rheum Dis* 2013. Dec 18. doi:10.1136/annrheumdis-2013-204436. [Epub ahead of print].
2. Said-Nahal R, Miceli-Richard C, Berthelot JM, Duché A, Dernis-Labous E, Blévec G Le, et al. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. *Groupe Français d'Etude Génétique des Spondylarthropathies. Arthritis Rheum* 2000;43:1356–1365.
3. Said-Nahal R, Miceli-Richard C, D'Agostino MA, Dernis-Labous E, Berthelot JM, Duché A, et al. Phenotypic diversity is not determined by independent genetic factors in familial spondylarthropathy. *Arthritis Rheum* 2001;45:478–484.
4. Breban M, Miceli-Richard C, Zinovieva E, Monnet D, Said-Nahal R. The genetics of spondylarthropathies. *Joint Bone Spine* 2006;73:355–362.
5. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–1337.
6. Maksymowych WP, Inman RD, Gladman DD, Reeve JP, Pope A, Rahman P. Association of a specific ERAP1/ARTS1 haplotype with disease susceptibility in ankylosing spondylitis. *Arthritis Rheum* 2009;60:1317–1323.
7. Tsui FWL, Haroon N, Reveille JD, Rahman P, Chiu B, Tsui HW, et al. Association of an ERAP1 ERAP2 haplotype with familial ankylosing spondylitis. *Ann Rheum Dis* 2010;69:733–736.
8. Davidson SI, Wu X, Liu Y, Wei M, Danoy PA, Thomas G, et al. Association of ERAP1, but not IL23R, with ankylosing spondylitis in a Han Chinese population. *Arthritis Rheum* 2009;60:3263–3268.
9. Kadi A, Izac B, Said-Nahal R, Leboime A, Praet L Van, Vlam K de, et al. Investigating the genetic association between ERAP1 and spondyloarthritis. *Ann Rheum Dis* 2013;72:608–613.

10. Cherciu M, Popa LO, Bojinca M, Dutescu MI, Bojinca V, Bara C, et al. Functional variants of ERAP1 gene are associated with HLA-B27 positive spondyloarthritis. *Tissue Antigens* 2013;82:192–196.
11. Evans DM, Spencer CCA, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–767.
12. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, Pointon JP, Robinson PC, Karaderi T, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* 2013;45:730–738.
13. Saric T, Chang S-C, Hattori A, York IA, Markant S, Rock KL, et al. An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat Immunol* 2002;3:1169–1176.
14. York IA, Chang S-C, Saric T, Keys JA, Favreau JM, Goldberg AL, et al. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat Immunol* 2002;3:1177–1184.
15. Cui X, Hawari F, Alsaaty S, Lawrence M, Combs CA, Geng W, et al. Identification of ARTS-1 as a novel TNFR1-binding protein that promotes TNFR1 ectodomain shedding. *J Clin Invest* 2002;110:515–526.
16. Cui X, Rouhani FN, Hawari F, Levine SJ. An aminopeptidase, ARTS-1, is required for interleukin-6 receptor shedding. *J Biol Chem* 2003;278:28677–28685.
17. Cui X, Rouhani FN, Hawari F, Levine SJ. Shedding of the type II IL-1 decoy receptor requires a multifunctional aminopeptidase, aminopeptidase regulator of TNF receptor type 1 shedding. *J Immunol* 2003;171:6814–6819.
18. Haroon N, Tsui FWL, Chiu B, Tsui HW, Inman RD. Serum cytokine receptors in ankylosing spondylitis: relationship to inflammatory markers and endoplasmic reticulum aminopeptidase polymorphisms. *J Rheumatol* 2010;37:1907–1910.
19. Harvey D, Pointon JJ, Evans DM, Karaderi T, Farrar C, Appleton LH, et al. Investigating the genetic association between ERAP1 and ankylosing spondylitis. *Hum Mol Genet* 2009;18:4204–4212.
20. Bettencourt BF, Rocha FL, Alves H, Amorim R, Caetano-Lopes J, Vieira-Sousa E, et al. Protective effect of an ERAP1 haplotype in ankylosing spondylitis: investigating non-MHC genes in HLA-B27-positive individuals. *Rheumatology (Oxford)* 2013;52:2168–2176.

21. Kochan G, Krojer T, Harvey D, Fischer R, Chen L, Vollmar M, et al. Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc Natl Acad Sci USA* 2011;108:7745–7750.
22. Goto Y, Hattori A, Ishii Y, Tsujimoto M. Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. *FEBS Lett* 2006;580:1833–1838.
23. Evnouchidou I, Kamal RP, Seregin SS, Goto Y, Tsujimoto M, Hattori A, et al. Cutting Edge: Coding single nucleotide polymorphisms of endoplasmic reticulum aminopeptidase 1 can affect antigenic peptide generation in vitro by influencing basic enzymatic properties of the enzyme. *J Immunol* 2011;186:1909–1913.
24. Reeves E, Edwards CJ, Elliott T, James E. Naturally Occurring ERAP1 Haplotypes Encode Functionally Distinct Alleles with Fine Substrate Specificity. *J Immunol* 2013;191:35–43.
25. Seregin SS, Rastall DPW, Evnouchidou I, Aylsworth CF, Quiroga D, Kamal RP, et al. Endoplasmic reticulum aminopeptidase-1 alleles associated with increased risk of ankylosing spondylitis reduce HLA-B27 mediated presentation of multiple antigens. *Autoimmunity* 2013.
26. Nguyen TT, Chang S-C, Evnouchidou I, York IA, Zikos C, Rock KL, et al. Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. *Nat Struct Mol Biol* 2011;18:604–613.
27. Hacquard-Bouder C, Falgarone G, Bosquet A, Smaoui F, Monnet D, Ittah M, et al. Defective costimulatory function is a striking feature of antigen-presenting cells in an HLA-B27-transgenic rat model of spondylarthropathy. *Arthritis Rheum* 2004;50:1624–1635.
28. Hacquard-Bouder C, Chimenti M-S, Giquel B, Donnadieu E, Fert I, Schmitt A, et al. Alteration of antigen-independent immunologic synapse formation between dendritic cells from HLA-B27-transgenic rats and CD4+ T cells: selective impairment of costimulatory molecule engagement by mature HLA-B27. *Arthritis Rheum* 2007;56:1478–1489.
29. Glatigny S, Fert I, Blaton MA, Lories RJ, Araujo LM, Chiocchia G, et al. Proinflammatory Th17 cells are expanded and induced by dendritic cells in spondylarthritis-prone HLA-B27-transgenic rats. *Arthritis Rheum* 2012;64:110–120.
30. Utriainen L, Firmin D, Wright P, Cerovic V, Breban M, McInnes I, et al. Expression of HLA-B27 causes loss of migratory dendritic cells in a rat model of spondylarthritis. *Arthritis Rheum* 2012;64:3199–3209.
31. Araujo LM, Fert I, Jouhault Q, Labroquère K, Andrieu M, Chiocchia G, et al. Increased production of interleukin-17 over interleukin-10 by treg cells implicates inducible

costimulator molecule in experimental spondyloarthritis. *Arthritis Rheumatol* 2014;66:2412–2422.

32. Talpin A, Costantino F, Bonilla N, Leboime A, Letourneur F, Jacques S, et al. Monocyte-derived dendritic cells from HLA-B27+ axial spondyloarthritis (SpA) patients display altered functional capacity and deregulated gene expression. *Arthritis Res Ther* 2014;16:417.

33. Rudwaleit M, Heijde D van der, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–783.

34. Rudwaleit M, Heijde D van der, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.

35. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265.

36. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;20:307–315.

37. Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, Edwards S, et al. Genetic inheritance of gene expression in human cell lines. *Am J Hum Genet* 2004;75:1094–1105.

38. Martín-Esteban A, Gómez-Molina P, Sanz-Bravo A, López de Castro JA. Combined effects of ankylosing spondylitis-associated ERAP1 polymorphisms outside the catalytic and peptide-binding sites on the processing of natural HLA-B27 ligands. *J Biol Chem* 2014;289:3978–3990.

39. Haroon N, Tsui FW, Uchanska-Ziegler B, Ziegler A, Inman RD. Endoplasmic reticulum aminopeptidase 1 (ERAP1) exhibits functionally significant interaction with HLA-B27 and relates to subtype specificity in ankylosing spondylitis. *Ann Rheum Dis* 2012;71:589–595.

40. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KCC, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202–1207.

41. Ding J, Gudjonsson JE, Liang L, Stuart PE, Li Y, Chen W, et al. Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. *Am J Hum Genet* 2010;87:779–789.

42. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* 2012;44:502–510.

43. Campbell EC, Fettke F, Bhat S, Morley KD, Powis SJ. Expression of MHC class I dimers and ERAP1 in an ankylosing spondylitis patient cohort. *Immunology* 2011;133:379–385.
44. Goto Y, Ogawa K, Hattori A, Tsujimoto M. Secretion of endoplasmic reticulum aminopeptidase 1 is involved in the activation of macrophages induced by lipopolysaccharide and interferon-gamma. *J Biol Chem* 2011;286:21906–21914.
45. Fruci D, Ferracuti S, Limongi MZ, Cunsolo V, Giorda E, Fraioli R, et al. Expression of endoplasmic reticulum aminopeptidases in EBV-B cell lines from healthy donors and in leukemia/lymphoma, carcinoma, and melanoma cell lines. *J Immunol* 2006;176:4869–4879.
46. García-Medel N, Sanz-Bravo A, Nguyen D Van, Galocha B, Gómez-Molina P, Martín-Esteban A, et al. Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 1 polymorphism and HLA-B27 in vivo. *Mol Cell Proteomics* 2012;11:1416–1429.
47. García-Medel N, Sanz-Bravo A, Alvarez-Navarro C, Gómez-Molina P, Barnea E, Marcilla M, et al. Peptide handling by HLA-B27 subtypes influences their biological behavior, association with ankylosing spondylitis and susceptibility to ERAP1. *Mol Cell Proteomics* 2014.
48. Akram A, Lin A, Gracey E, Streutker CJ, Inman RD. HLA-B27, but not HLA-B7, immunodominance to influenza is ERAP dependent. *J Immunol* 2014;192:5520–5528.
49. Marcilla M, López de Castro JA. Peptides: the cornerstone of HLA-B27 biology and pathogenetic role in spondyloarthritis. *Tissue Antigens* 2008;71:495–506.



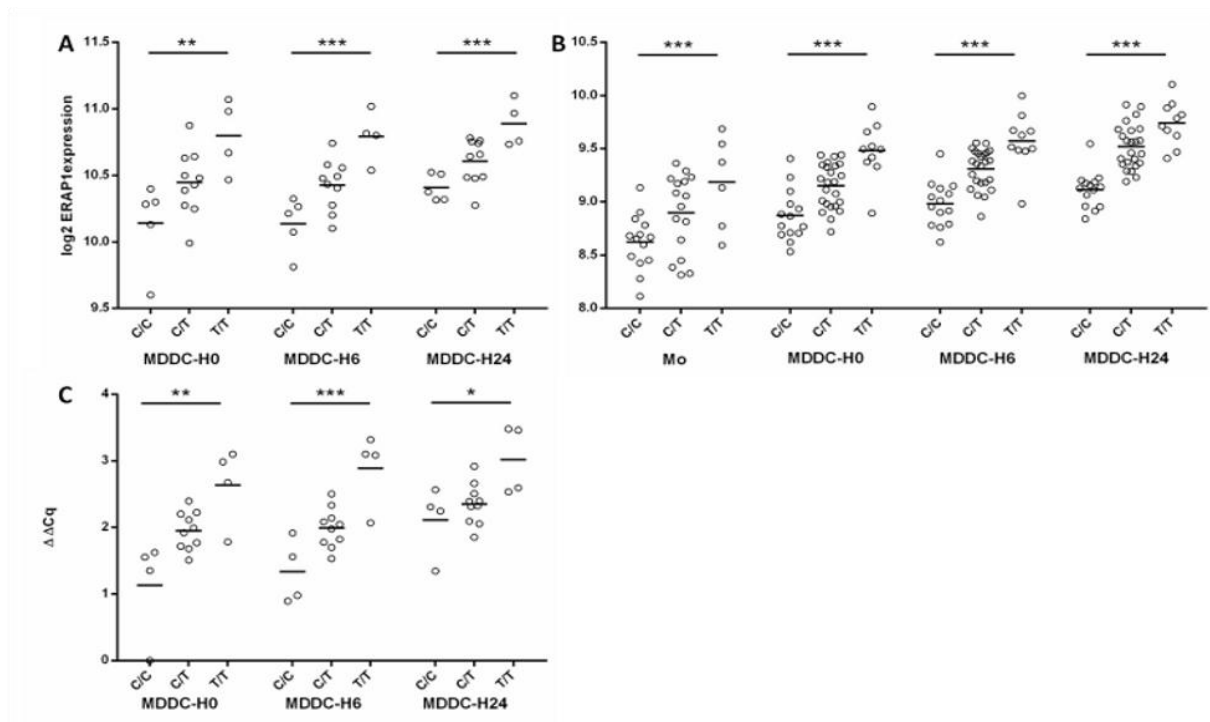
## TABLES

**Table 1.** Clinical characteristics of SpA patients in the study

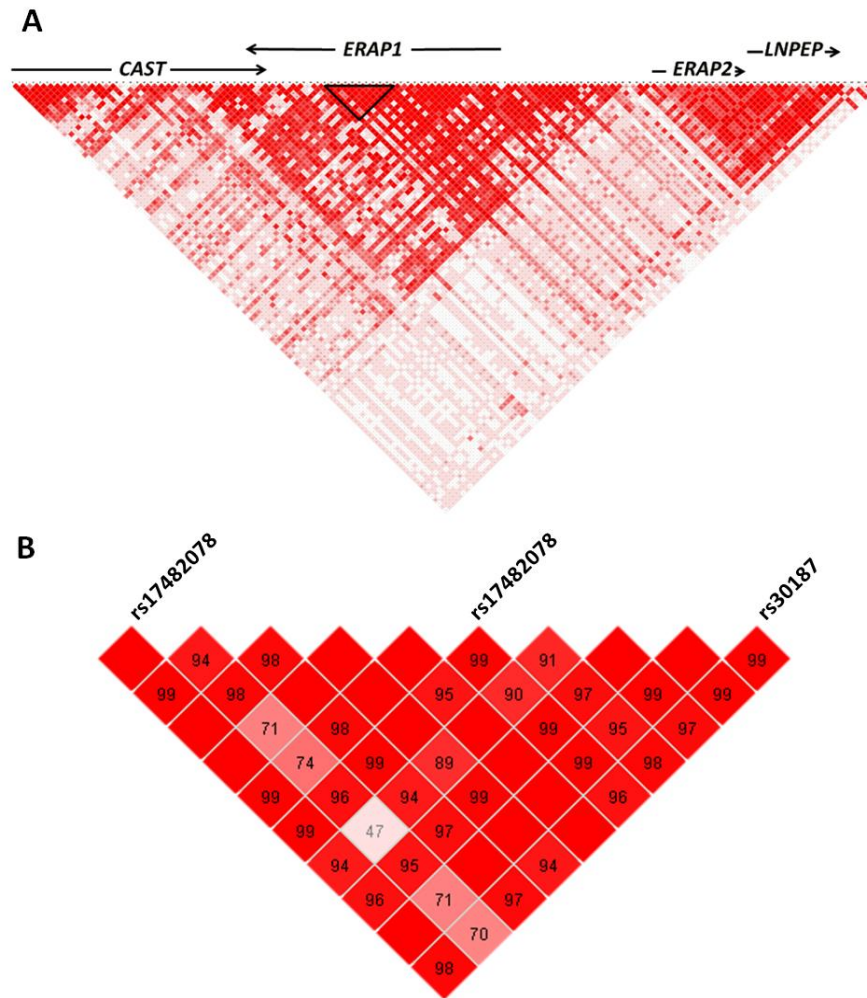
Characteristic	Discovery (n = 9)	Replication (n = 14)	B-LCL (n = 7**)	AS Immunochip (n = 354)
Age in years, mean $\pm$ SD	48.1 $\pm$ 7.9	47.1 $\pm$ 7.6	48.1 $\pm$ 16.2	48.1 $\pm$ 14.2
Disease duration in years, mean $\pm$ SD	15.1 $\pm$ 5.2	22.4 $\pm$ 7.3	21.2 $\pm$ 15.0	24.5 $\pm$ 13.2
Sex ratio, men:women (% of men)	5:4 (55.6)	8:6 (57.1)	4:3 (57.1)	219:135 (61.9)
<i>HLA-B27</i> positivity, %	100	100	100	91.2
Axial manifestations				
- Inflammatory back pain, %	100	100	100	100
- Radiographic sacroiliitis, %*	89	64.3	100	100
Peripheral manifestations				
- Peripheral arthritis, %	22.2	42.9	66.7	45.3
- Peripheral enthesitis, %	66.7	85.7	50	73.7
Extra-articular manifestations				
- Uveitis, %	33.3	28.6	66.7	41.1
- Psoriasis, %	33.3	7.1	0	26.0
- Inflammatory bowel disease, %	0	0	0	8.6
Current treatment				
- NSAID, %	55.6	75.0***	NA	NA
- Anti-TNF therapy, %	22.2	0***	NA	NA
- None, %	22.2	25.0***	NA	NA

The registered manifestations correspond to those present at the time of examination, or retrieved from past-medical history. AS: ankylosing spondylitis; B-LCL: B -Lymphoblastoid cell lines; SD: standard deviation; Inflammatory bowel disease: Crohn's disease or ulcerative colitis; NSAID: Non-steroidal anti-inflammatory drug; NA: not available; \* Refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral; \*\* clinical data available for 6 patients; \*\*\* treatment data available for 12 patients.

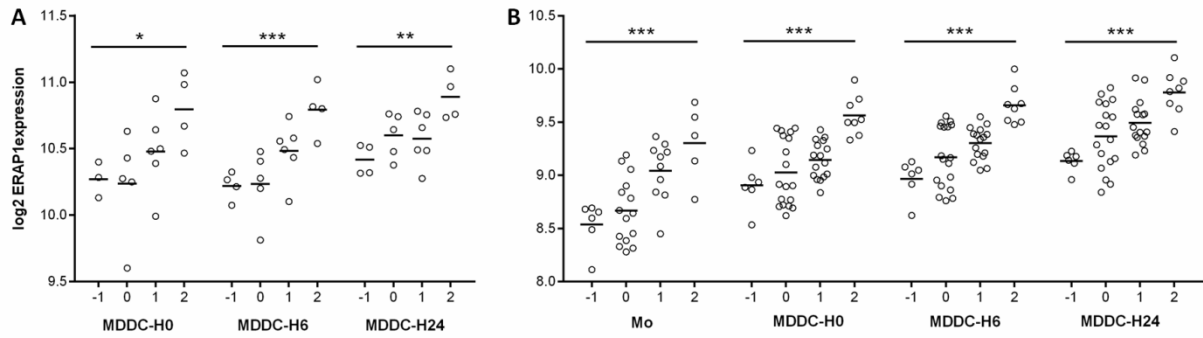
## FIGURES



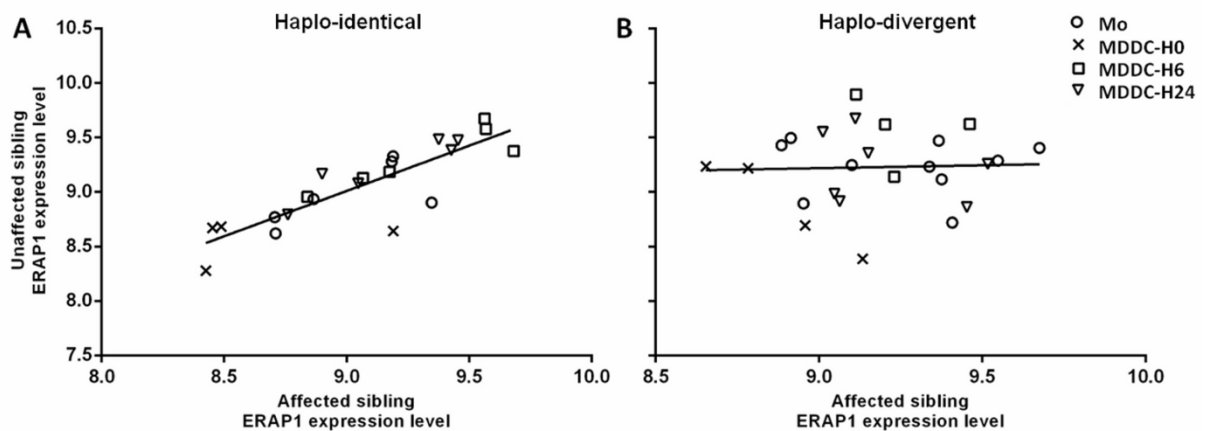
**Figure 1. *ERAP1* gene expression is correlated to the rs30187 genotype in MD-DCs and circulating monocytes.** *ERAP1* mRNA levels were evaluated by micro-array in the discovery (A) and replication (B) cohorts and by quantitative real-time PCR (C) in the discovery samples. The x axis indicates the rs30187 genotype of individuals in the cohorts (C/C, C/T, or T/T) and the cell condition: (A,B,C) MD-DCs at different stimulation time-points (H0, H6, H24) or (B) monocytes (Mo) and the y axis shows the log2 of *ERAP1* fluorescence intensity as a measure of gene expression (A,B) or  $\Delta\Delta Cq$  (C). \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .



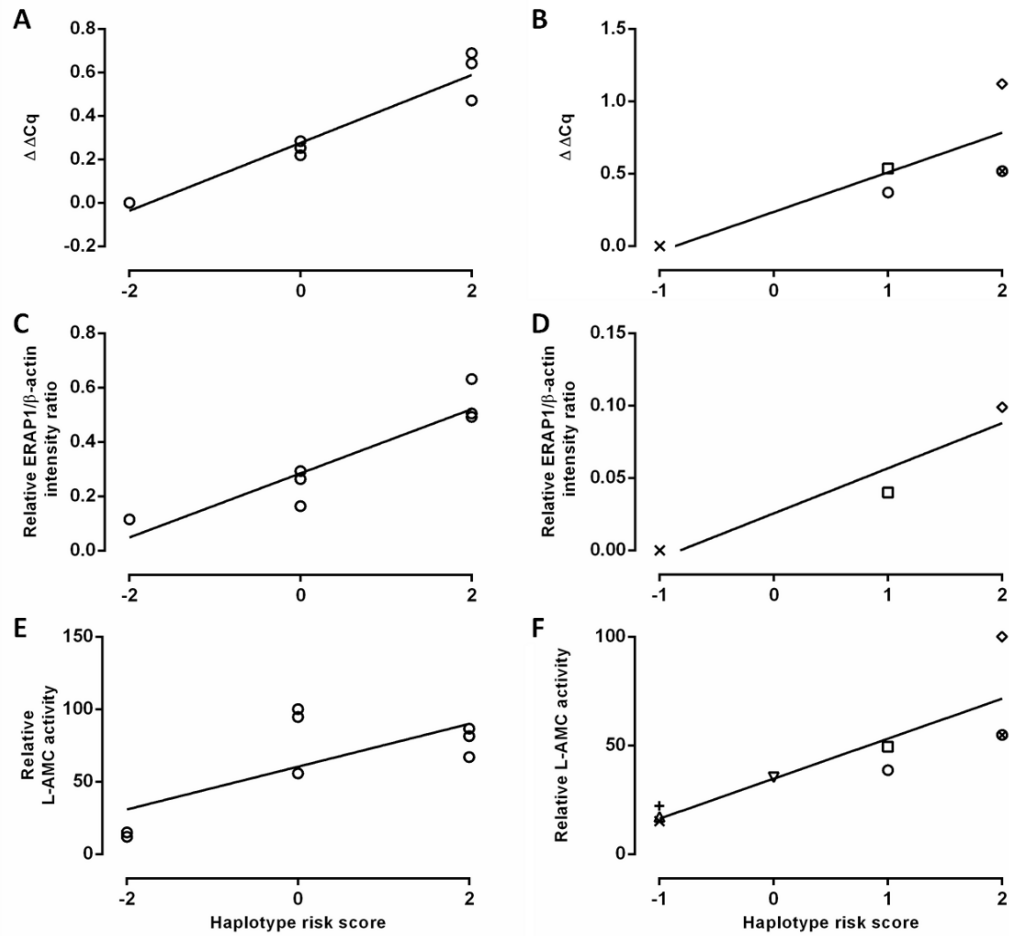
**Figure 2. Haploview linkage disequilibrium (LD) plot covering the region surrounding *ERAP1* (A), including the rs17482078/rs10050860/rs30187 haplotype (black triangle in (A), higher magnification in (B)). The darker the box, the stronger the LD. The number inside boxes in (B) represents  $|D'|$ .**



**Figure 3. *ERAP1* gene expression assessed by micro-arrays, according to haplotype risk score in discovery (A) and replication (B) cohorts.** The x axis of the plots represents the haplotype risk score (from -1 to 2) according to the cell condition: (A,B) MD-DC stimulation time-point (MDDC-H0, MDDC-H6, MDDC-H24) or (B) monocytes (Mo) and the y axis shows the log2 of *ERAP1* fluorescence intensity as a measure of gene expression. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .



**Figure 4. Correlation of *ERAP1* expression levels between affected and unaffected siblings in haplo-identical (A) or haplo-divergent (B) pairs.** The x axis of the plots represents the log2 of *ERAP1* fluorescence intensity (as a measure of gene expression assessed by micro-arrays) of the affected sibling for each sibling pair and for each cell condition (Mo: monocytes and MD-DCs) and stimulation time-point (H0, H6, H24) and the y axis shows *ERAP1* gene expression of the corresponding unaffected sibling.



**Figure 5. Correlation between *ERAP1* haplotype risk score and gene expression, protein expression and enzymatic activity in B-LCLs (A,C,E) and in MD-DCs (B,D,F)** The x axis of the plots represents the haplotype risk score and the y axis shows (A,B) *ERAP1* gene expression assessed by  $\Delta\Delta Cq$ , (C,D) *ERAP1* protein level assessed by Western-blot as *ERAP1*/beta-actin intensity ratio and (E,F) *ERAP1* enzymatic activity assessed on L-AMC hydrolysis.

**Supplementary table 1. Distribution of *ERAPI* rs17492078/rs10050860/rs30187 haplotype risk score in SpA patients and controls from the French Immunochip cohort.**

Haplotype risk score	Controls (n = 346 )	Patients (n = 354)
-2	13	9
-1	70	41
0	116	125
1	107	121
2	40	58

Each individual from the French Immunochip cohort was attributed a score corresponding to its rs17492078/rs10050860/rs30187 *ERAPI* haplotypes combination, according to the following additive rule: each CCT (susceptibility haplotype) counted +1, each CCC (neutral haplotype) counted 0, and each TTC (protective haplotype) counted -1. The distribution was significantly different between patients and controls, validating this genotype score as associated with disease risk (linear regression analysis;  $p = 0.002$  ; odds ratio = 1.27; 95% confidence interval: 1.1 to 1.48)

**TROISIÈME PARTIE :**  
**DISCUSSION & PERSPECTIVES**

## ***Discussion et perspectives***

---

Comme exposé précédemment, mon travail de thèse s'est organisé autour de deux axes distincts, mais complémentaires, de génétique et de génomique fonctionnelle. La discussion des résultats obtenus s'articulera donc autour de ces deux axes.

### **Études familiales : un outil pour l'identification de gènes de susceptibilité dans la spondyloarthrite**

Dans la première partie de ce travail de thèse, des familles multiplex de SpA ont été étudiées. Les données de génotypage obtenues grâce à des puces à SNP de haute densité couvrant tout le génome ont permis de réaliser à la fois une étude de liaison et une étude d'association intra-familiale.

### **Positionnement du sujet et problématique générale**

L'étude génétique des maladies s'est traditionnellement concentrée sur les maladies mendéliennes, c'est-à-dire présentant une hérédité facilement modélisable, dans laquelle un gène unique peut être rendu responsable du phénotype. Les analyses de liaison et le clonage positionnel ont permis l'identification d'un grand nombre de gènes responsables de ce type de pathologie. Elles ont ainsi abouti à des avancées notables dans leur diagnostic et leur dépistage<sup>416</sup>. Plus récemment, il a été démontré que des pathologies fréquentes, non héréditaires à proprement parler, telles que les maladies cardiovasculaires ou les maladies inflammatoires, possédaient elles-aussi une composante génétique<sup>417</sup>. Cette composante est toutefois beaucoup plus complexe que dans les maladies mendéliennes, puisque la présence simultanée de nombreuses variations génétiques au sein de différents gènes en association avec des facteurs environnementaux est nécessaire à la survenue de la maladie. Ces pathologies sont dites complexes ou multifactorielles.

Parmi ces maladies multifactorielles, la SpA peut être considérée comme particulière du fait de l'importance de sa composante génétique. En effet, l'indice de récurrence au premier degré a été évalué à 40 pour la SpA dans son ensemble et à 80 pour la SA<sup>222,231</sup>. À titre de comparaison, cet indice est d'environ 15 pour le diabète de type 1, 8 pour la schizophrénie et 3,5 pour le diabète de type 2<sup>417</sup>. Cette composante génétique est principalement due à l'allèle HLA-B27 dont la forte association à la SpA a été identifiée dès le début des années 1970<sup>98,99</sup>. Il a toutefois fallu attendre l'avènement des GWAS pour mettre en évidence de façon certaine d'autres facteurs génétiques de susceptibilité<sup>256,267-270</sup>. À ce jour, plus d'une vingtaine d'entre eux ont été identifiés. Cependant,



l'ensemble de ces locus est loin d'expliquer l'ensemble de la prédisposition à la maladie<sup>256</sup>. Dès lors, se pose la question de la manière d'identifier les autres facteurs génétiques en cause.

Depuis de nombreuses années, mon équipe d'accueil a fait le choix de privilégier les études familiales aux études cas-témoins. L'inclusion de familles multiplex de SpA s'est faite via le GFEGS créé en 1995 sous l'impulsion du Pr Breban. Ce groupe est un réseau de rhumatologues hospitaliers participant au recrutement de formes multiplex de SpA. L'inclusion des familles est centralisée par un médecin rhumatologue, le Dr Said-Nahal, chargé de recueillir les données démographiques, cliniques et radiographiques des patients après avoir vérifié leur diagnostic. L'un des avantages des études familiales par rapport aux études cas-témoins est qu'elles sont plus puissantes pour l'identification des variants rares<sup>418,419</sup>, dont on suppose qu'ils participent substantiellement à l'hérédité des maladies complexes<sup>420</sup>. Elles permettent aussi de s'affranchir du biais de stratification et de détecter de façon plus efficace les erreurs de génotypage<sup>421</sup>.

Une autre particularité de notre équipe est le choix du phénotype étudié (la SpA dans son ensemble), alors que la plupart des autres équipes se sont restreintes aux cas de SA. Le fait de n'étudier que la SA, qui requiert la présence d'une sacro-iliite radiologique évoluée, permet d'obtenir une population plus homogène et évite d'analyser ensemble des affections au déterminisme génétique partagé seulement en partie. Toutefois, de nombreux éléments plaident pour retenir la SpA au sens large comme phénotype d'étude, tels que la coexistence de plusieurs sous-types de SpA au sein d'une même famille<sup>97,228,229,422</sup> ou la présence de facteurs génétiques communs à l'ensemble des sous-types, tels que HLA-B27 ou *ERAP1*<sup>100,383</sup>. Ces éléments suggèrent l'existence d'un terrain génétique commun à l'ensemble des SpA auquel peuvent s'associer d'autres facteurs génétiques modulant le phénotype de la maladie.

## Résultats préliminaires et objectifs du projet

En dehors des analyses phénotypiques familiales décrites précédemment, le GFEGS a entrepris dans les années 2000 une première étude de liaison pangénomique. Cette étude, basée sur le génotypage de microsatellites, a abouti à la découverte d'un locus significativement lié en 9q31-34<sup>260</sup>, présentant déjà un signal de liaison dans une cohorte britannique<sup>257,258</sup>. Cette région, dénommée SPA2, a fait l'objet d'une étude approfondie au sein de notre laboratoire afin d'identifier le ou les facteurs de susceptibilité à l'origine du signal de liaison<sup>264,266</sup>. L'intervalle d'intérêt a ainsi pu être réduit à un haplotype de six SNPs couvrant une région de 72 kb dont le séquençage est en cours d'analyse.

L'une des limites des analyses de liaison basées sur les microsatellites est leur faible densité en marqueurs, conduisant à une imprécision sur la localisation des pics de liaison<sup>423</sup>. Ainsi, dans la plupart de ces études, les régions de liaison couvrent plusieurs Mb et l'identification du ou des facteurs de susceptibilité à l'origine du signal de liaison est souvent ardue<sup>424</sup>. Pour tenter de remédier à cet inconvénient, nous avons entrepris de génotyper un nouveau jeu plus étendu de familles multiplex de SpA (dont certaines avaient cependant déjà été génotypées dans la précédente étude de liaison) à l'aide de puces Affymetrix 250K permettant le génotypage de 260 000 marqueurs répartis sur l'ensemble du génome. À l'aide de ces données, nous avons ensuite effectué des analyses de liaison et d'association intra-familiale dont les résultats, présentés dans les **articles 1 et 2**, sont discutés ici.

### Analyse de liaison

Après un contrôle qualité rigoureux, nous avons d'abord réalisé une analyse de liaison non paramétrique sur l'ensemble du génome. Comme attendu, la région la plus significativement liée à la SpA s'est révélée être la région du CMH comprenant l'allèle HLA-B27. En dehors de cette région, un seul locus était significativement lié à la SpA sur le bras long du chromosome 13 (13q13). Par ailleurs, huit autres locus, dont la région 9q33 correspondant à SPA2, obtenaient un LOD score suggestif.

### Locus 13q13

La liaison significative du locus 13q13 à la SpA est rapportée pour la première fois. Il faut toutefois noter la présence d'un signal suggestif dans la précédente étude du GFECS basée sur les microsatellites et dans la méta-analyse des trois études de liaison publiées dans la SpA réalisée par Carter *et al.*<sup>260,262</sup>. Par ailleurs, une liaison significative entre cette région et la maladie de Crohn, fréquemment associée à la SpA, a été rapportée en 2008<sup>404</sup>.

L'étude approfondie du locus 13q13 a permis d'affiner la région d'intérêt à un intervalle d'environ 1 Mb situé entre 38,753 et 40,040 Mb. Si peu de données sont disponibles pour beaucoup des gènes contenus dans cet intervalle, certains se révèlent de bons gènes candidats, du fait de leur fonction (comme *UFM1* impliqué dans les phénomènes de stress du réticulum endoplasmique)<sup>425,426</sup> ou parce que certains de leurs polymorphismes ont été associés à des pathologies proches de la SpA (comme *LHFP* et le psoriasis)<sup>427</sup>.

### Complexe majeur d'histocompatibilité

Le CMH est donc la région qui a obtenu le LOD score le plus élevé. Plus de 92% des patients étaient HLA-B27+, soulignant le poids de ce facteur génétique dans la susceptibilité à la maladie.

Toutefois, une hétérogénéité allélique a été observée dans trois familles où aucun des individus malades n'était porteur de l'allèle B27, mais où tous portaient un haplotype commun dans la région du CMH. Pour l'une de ces familles, nous disposions du résultat du typage HLA-B et nous avons donc pu déterminer que l'allèle HLA-B62 ségrégeait avec la maladie.

Les reconstructions haplotypiques dans la région du CMH ont permis de constater que la grande majorité des familles étaient liées à ce locus. Plusieurs raisons expliquaient l'absence de liaison observée pour certaines familles. Dans de nombreux cas, il existait une association sans liaison en raison de la présence de deux haplotypes porteurs de l'allèle B27 provenant de plusieurs individus fondateurs. Dans d'autres familles, HLA-B27 semblait être en cause dans l'agrégation, mais la présence d'un ou plusieurs sujets atteints HLA-B27- indiquait la participation importante d'autres facteurs de susceptibilité et expliquait l'absence de liaison observée. Enfin dans une seule famille, ni le HLA-B27 ni le CMH ne semblaient expliquer l'agrégation, puisqu'aucun des malades n'était HLA-B27+ et qu'ils ne partageaient aucun haplotype du CMH.

### Comparaison avec les résultats de la littérature

En dehors du CMH, il y a peu de concordance entre nos résultats et ceux des études de liaison déjà publiées<sup>258-260</sup>. Trois hypothèses peuvent être avancées pour expliquer ces différences : une hétérogénéité génétique, une hétérogénéité phénotypique ou des différences méthodologiques.

La plupart des maladies complexes impliquent un grand nombre de gènes de susceptibilité. Une étude récente de Joshi *et al.* suggère que la composante génétique des formes familiales de SpA diffère peu de celle des formes sporadiques en dehors d'une fréquence plus élevée du HLA-B27<sup>428</sup>. Il est alors probable que les gènes en cause diffèrent en fonction des familles. Cette hypothèse se confirme dans notre étude où les analyses paramétriques ont permis d'estimer qu'environ 30% des familles étaient liées au locus 13q13 et 15% en 9q33. Cette hétérogénéité génétique entre les différentes études de liaison est d'autant plus probable que l'origine géographique des patients était différente<sup>429</sup>.

L'hypothèse d'une hétérogénéité phénotypique peut également être évoquée. Comme déjà souligné, nous avons pris le parti d'étudier la SpA dans son ensemble, tandis que les autres études s'étaient restreintes aux cas de SA. Si l'on compare les caractéristiques cliniques des patients inclus dans notre étude avec celles des autres, la principale différence était la moindre fréquence de sacro-iliite radiologique chez nos malades (59% dans notre cohorte). Parmi les autres différences significatives, on observait un ratio homme:femme plus faible, une fréquence plus faible d'uvéite, de MICI et une fréquence plus importante de psoriasis cutané. Plusieurs études suggèrent que les

atteintes structurales de la SpA sont en partie génétiquement déterminées<sup>255,384,430</sup>. Par conséquent, certains des locus identifiés dans les études restreintes à la SA pourraient correspondre à des facteurs de sévérité plutôt qu'à des facteurs de susceptibilité à la maladie.

La dernière raison pouvant expliquer ces différences est liée à la méthodologie utilisée. Les trois précédentes analyses de liaison étaient basées sur l'utilisation d'un faible nombre de marqueurs microsatellites (<400), tandis que nous avons utilisé des polymorphismes bi-alléliques avec une répartition très dense sur tout le génome. Pris individuellement, les SNPs sont moins informatifs que les microsatellites car ils ne possèdent que deux allèles. Ce manque d'informativité est compensé par la possibilité de les génotyper en très grand nombre. Il a ainsi été démontré qu'une carte dense de SNPs permettait d'obtenir une meilleure informativité qu'une carte peu dense de microsatellites<sup>431</sup>. Dans certaines études de liaison, le même jeu de familles a été génotypé à la fois à l'aide de microsatellites et de SNPs. Les résultats obtenus différaient en fonction du type de marqueur utilisé à la fois sur le nombre de locus identifiés, le niveau d'intensité de la liaison et la localisation du pic de liaison, le plus souvent à l'avantage de l'analyse basée sur les SNPs<sup>432-436</sup>.

### Justification de la méthodologie et difficultés rencontrées

L'approche expérimentale proposée est originale à plusieurs titres. Tout d'abord, les études de liaison sont devenues moins courantes que les études d'association depuis plusieurs années. De plus, compte-tenu du faible nombre d'analyses de liaison réalisées récemment, peu d'entre elles ont bénéficié des progrès des méthodes de génotypage et l'utilisation de marqueurs bi-alléliques de haute densité est donc peu fréquente dans ce contexte.

Le choix de l'analyse de liaison rejoint celui des études familiales discuté précédemment. L'objectif est d'identifier des variants rares, peu accessibles aux GWAS. Il a été reproché aux études de liaison d'être peu précises dans la localisation des régions d'intérêt<sup>423</sup> et de rarement aboutir à l'identification des facteurs à l'origine de la liaison. Une manière d'améliorer cette identification est l'utilisation conjointe de techniques de liaison et de techniques d'association, que ce soit dans des études d'association « classiques » ou dans des études faisant intervenir le séquençage dit « de nouvelle génération » (next generation sequencing ou NGS)<sup>437,438</sup>.

L'une des difficultés rencontrées dans les études utilisant le NGS pour identifier des facteurs de susceptibilité génétiques dans les maladies complexes est le grand nombre de variants mis en évidence<sup>439</sup>. La combinaison de données de liaison et de séquençage haut-débit peut permettre de mieux guider la sélection des variants et d'identifier les polymorphismes responsables du signal de liaison<sup>438</sup>. Ainsi, dans une étude cherchant à identifier les déterminants génétiques de l'homéostasie,

du glucose et de l'adiposité, Bowden *et al.* ont mis en évidence dans certaines familles une forte liaison entre le locus 3q27 et le taux sérique d'adiponectine. Le séquençage de l'exome d'individus appartenant à ces familles a permis l'identification d'un variant rare à l'intérieur du gène ADIPOQ contribuant à la variation du taux sérique d'adiponectine et expliquant le pic de liaison<sup>440</sup>. Des stratégies similaires ont été appliquées avec succès à d'autres traits complexes comme la cardiomyopathie dilatée<sup>441</sup>, l'autisme<sup>442</sup> ou le psoriasis<sup>443</sup>.

Le choix du type de marqueurs utilisés a aussi été discuté précédemment. Notre objectif en utilisant des puces à SNP de haute densité était d'obtenir une puissance de détection plus importante et d'améliorer la caractérisation des facteurs en cause à l'intérieur des régions liées. Nous avons ainsi pu identifier une liaison significative en 13q13, alors que ce locus n'avait obtenu qu'un score de liaison suggestif dans la précédente analyse de liaison ayant utilisé des microsatellites et ayant 47 familles en commun avec l'étude actuelle.

L'utilisation de puces à SNPs de haute densité a aussi été à l'origine de difficultés méthodologiques. En effet, la plupart des logiciels d'analyse de liaison limitent le nombre de marqueurs pouvant être analysés simultanément. Nous avons opté pour le logiciel Merlin qui permet une grande flexibilité au niveau du nombre de marqueurs utilisés comme au niveau de la complexité des pédigrés analysés. Cette souplesse se fait cependant au prix d'un besoin important en ressources informatiques et d'un temps d'analyse extrêmement long.

Un autre point important à considérer dans les analyses de liaison à partir de SNPs est de s'assurer de la qualité de la carte utilisée. Les informations fournies par Affymetrix concernant la puce Affy250K se sont en effet révélées en partie obsolètes. La mise à jour des informations en fonction des données de dbSNP135 a conduit au retrait de 5587 SNPs de l'analyse et à l'actualisation des positions. Cette étape, comme celle du contrôle qualité, est cruciale car des erreurs de localisation ou de génotypage peuvent influencer de manière importante sur les résultats et notamment créer de faux signaux de liaison<sup>444–447</sup>.

### Perspectives

Les résultats obtenus grâce à cette étude de liaison nécessitent d'être approfondis. Pour ce faire, nous avons déjà génotypé un nouveau jeu de 50 familles multiplex à l'aide des mêmes puces Affy250K. Une analyse de liaison de ces familles et des deux jeux de familles fusionnés est en cours afin de répliquer la liaison observée sur le chromosome 13 et d'identifier de nouveaux locus liés à la maladie, notamment parmi les huit présentant un score de liaison suggestif.

Le résultat le plus intéressant de notre étude est l'identification d'un nouveau locus significativement lié à la SpA en 13q13. L'objectif est désormais d'identifier le (ou les) variant(s) de susceptibilité à l'origine de la liaison au sein de ce locus. Pour ce faire, l'intervalle de 1,3 Mb que nous avons délimité sera séquencé par une technique de NGS.

Pour chacune des familles dans lesquelles un haplotype unique co-ségrège avec la maladie, nous avons sélectionné un individu malade porteur de cet haplotype dont l'ADN sera séquencé. Dans la première phase de capture, l'ADN sera hybridé à des sondes couvrant la région d'intérêt. Cette étape sera suivie du séquençage des librairies enrichies. L'analyse des séquences permettra d'identifier puis de filtrer les variants candidats, par comparaison avec les séquences disponibles dans les bases de données publiques. Ces variants seront génotypés chez les autres individus de la famille afin de vérifier leur co-ségrégation à la maladie. La confirmation de l'implication des variants dans la SpA se fera ensuite à la fois par des études d'association dans des cohortes indépendantes et par des études fonctionnelles visant à déterminer l'impact de ces variants.

À l'aide des données de liaison collectées, il sera également intéressant de modéliser plus finement la transmission de la maladie. Au vu des premiers résultats obtenus, il semble qu'un modèle bi-locus puisse être envisagé pour les familles liées au chromosome 13, impliquant à la fois HLA-B27 et la région 13q13. De nouvelles analyses de liaison paramétriques sur le chromosome 13 intégrant différentes classes de risque en fonction du portage du HLA-B27 devraient nous permettre d'affiner notre modèle.

Enfin, nous envisageons d'étudier individuellement les plus grandes familles multiplex. Dans chacune de ces familles, nous projetons de séquencer l'exome de deux individus malades sélectionnés pour être aussi distants que possible dans l'arbre familial, de façon à limiter le nombre de variants partagés du fait du seul hasard. Les informations de liaison disponibles dans ces familles nous aideront à guider la sélection des variants candidats dont l'étude sera poursuivie.

### Analyse d'association

Dans les suites de l'analyse de liaison, nous avons réalisé une analyse d'association intra-familiale pangénomique à partir des mêmes données de génotypage (**Article 4**). En dehors du CMH, le score d'association le plus significatif concernait un polymorphisme intergénique situé sur le chromosome 12 (rs4765413), qui n'atteignait cependant pas le seuil significatif de niveau pangénomique, en appliquant une correction stringente de type Bonferroni. Néanmoins, à l'issue de deux études d'extension, basées sur l'analyse d'un deuxième jeu de trios et familles multiplex provenant de notre collection puis d'une cohorte familiale nord-américaine provenant d'une

collaboration avec le Pr. John Reveille (Université du Texas), nous avons pu mettre en évidence l'association à la SpA d'un polymorphisme situé au sein du gène *MAPK14* avec un score d'association proche de la significativité pangénomique.

### *MAPK14*

Les résultats les plus intéressants de cette étude concernent donc deux polymorphismes introniques du gène *MAPK14* dont l'association à la SpA est proche de la significativité. Ce gène, « mitogen-activated protein kinase 14 » code pour un membre de la famille des p38 MAP kinases qui sont d'importantes enzymes régulatrices de la synthèse de cytokine pro-inflammatoires, notamment du TNF $\alpha$ <sup>448</sup>. Il a été impliqué dans les MICI<sup>406</sup> et l'inhibition de cette voie a été envisagée comme possible cible thérapeutique<sup>449</sup>. Certains polymorphismes situés au sein de ce gène ou sur la voie de signalisation des p38 MAP kinases ont été associés à la réponse aux anti-TNF $\alpha$  dans la PR<sup>450</sup>. Enfin il a été démontré que la voie p38 MAPK était dérégulée dans des lignées monocytaires exprimant HLA-B27<sup>451,452</sup>. L'ensemble de ces éléments font de *MAPK14* un excellent gène candidat en ce qui concerne la susceptibilité à la SpA.

### *Comparaison avec les résultats de la littérature*

Dans notre étude, aucun des SNPs les plus fortement associés à la SpA n'est situé dans l'un des locus précédemment associés à la SA dans les GWAS. Trois explications peuvent être avancées : tout d'abord, la puce Affy 250K que nous avons utilisée contient des tag-SNPs destinés à couvrir au mieux l'ensemble du génome, mais ne contient que très peu de SNPs situés dans les régions codantes. Or, la plupart des GWAS ont été réalisés à partir d'un jeu de marqueurs plus dense<sup>268-270</sup> ou se sont concentrés sur les régions codantes<sup>256,267</sup>. De plus, le phénotype étudié n'est pas strictement identique puisque nous nous sommes intéressés à l'ensemble des SpA, alors que les GWAS se sont focalisés sur l'étude des SA. Or l'étude du polymorphisme rs11209026 de l'*IL23R* nous a permis de démontrer qu'il existait des facteurs spécifiques à certains sous-types de la maladie<sup>384</sup>. Enfin, ces différences peuvent refléter des différences de méthodologie : en effet, notre étude est basée sur un test de déséquilibre de transmission qui teste à la fois la liaison et l'association, tandis que les méthodes utilisées dans les études cas-témoins ne testent que l'association. Nous avons toutefois observé des associations nominales ( $p < 0,05$ ) entre des polymorphismes situés à proximité de facteurs de susceptibilité identifiés dans des GWAS (*IL23R*, *PTGER4*, *ERAP1*, *BACH2*, *ZMIZ1*) et la SpA.

Il est également intéressant d'observer que les SNPs les plus fortement associés à la SpA ne se situent pas dans les régions les plus liées à la maladie, alors que le jeu de familles utilisé est quasiment identique. Ce résultat peut s'expliquer par le fait qu'un nombre restreint de familles s'est

avéré lié à chacun des locus mis en évidence (par exemple, 30% des familles seulement pour le locus 13q13, le plus lié à la maladie en dehors du CMH). Dans ce contexte, pour un locus donné, il paraît intéressant de tester l'association des marqueurs de cette région dans le sous-groupe des familles les plus liées. Cependant, le gain potentiel en informativité des familles sélectionnées sur le critère de liaison est contre-balancé par une perte de puissance statistique liée au plus faible effectif.

### Perspectives

Les résultats obtenus concernant *MAPK14* nécessiteront d'être approfondis. Tout d'abord, une étude complémentaire de type cas-témoins sera réalisée afin de consolider le niveau de preuve de l'association entre des polymorphismes de ce gène et la SpA. Il est notamment prévu d'inclure dans cette étude un nombre important de témoins HLA-B27 positifs pour permettre de s'affranchir, autant que faire se peut, du déséquilibre de liaison qui pourrait exister entre ce gène et l'allèle HLA-B27 porté par le même chromosome. En effet, bien que ces deux locus soient séparés de près de 6 Mb, nous ne pouvons pas formellement écarter sur la base de nos données actuelles que le signal observé avec *MAPK14* ne soit le reflet de l'association entre HLA-B27 et SpA. Enfin, les deux polymorphismes significatifs identifiés étant introniques, il nous paraît intéressant de séquencer l'ensemble de ce gène afin d'identifier le(s) polymorphisme(s) causal(aux) pouvant expliquer l'association.

### De la maladie au sous-phénotype : exemple du gène *IL23R*

Si l'étude du phénotype SpA au sens large plutôt que celle restreinte à la SA se justifie sur le plan génétique, il est probable que l'hétérogénéité de présentation clinique des patients soit au moins en partie liée à des déterminants génétiques. Un exemple de polymorphisme modulant le phénotype de la maladie est montré dans l'**Article 3**. En effet, le polymorphisme non synonyme de l'*IL23R* rs11209026 n'est associé qu'aux formes de SpA avec une sacro-iliite radiologique.

### Facteurs génétiques associés à l'atteinte radiologique

Plusieurs études ont cherché à identifier les déterminants génétiques de l'atteinte structurale de la maladie. En 2009, Ward *et al.* ont montré que la sévérité radiographique était déterminée entre autres par l'âge de début de la maladie, le sexe, le tabagisme et la présence de certaines allèles du CMH (HLA-B\*4100, DRB1\*0804, DQA1\*0401, DQB1\*0603, DRB1\*0801, et DPB1\*0202)<sup>430</sup>. Nous avons nous-mêmes établi récemment que la présence du HLA-B27 était associée de manière indépendante à la présence d'une sacro-iliite radiologique selon les critères de New York modifiés (**Annexe 2**)<sup>453</sup>. Des polymorphismes en dehors du CMH ont également été



associés à la sévérité radiologique, notamment au sein des gènes *LMP2*, *RANK*, *PTGS1*, *ADRB1*, *NELL1* et *ERAP1*<sup>352,454–456</sup>.

### *Facteurs génétiques associés à l'activité et à la sévérité fonctionnelle*

D'autres études enfin se sont intéressées aux facteurs génétiques associés à l'activité (BASDAI) ou à la sévérité fonctionnelle (BASFI) de la maladie. Il a été montré en premier lieu que le BASDAI et le BASFI étaient dépendants d'une forte héritabilité<sup>254,255</sup>. À partir des données familiales utilisées pour les analyses de liaison du groupe d'Oxford, une sous-analyse a été réalisée pour tester s'il existait aussi une liaison avec plusieurs indices permettant d'apprécier la sévérité de la SpA (l'âge de début, le BASDAI et le BASFI). Une seule liaison significative a été mise en évidence entre le bras court du chromosome 18 et le BASDAI<sup>457</sup>. Aucune liaison n'a été détectée avec le CMH. D'autres études génétiques se sont par ailleurs intéressées à la sévérité de la maladie. Il s'agissait d'études d'association de type gène-candidat dont les résultats se sont souvent révélés contradictoires<sup>458–464</sup>.

## **Génomique fonctionnelle**

Dans la deuxième partie de ce travail, nous nous sommes intéressés au transcriptome des DCs dans la SpA. Nous avons en premier lieu cherché à identifier un profil d'expression génique propre aux DCs de patients atteints de SpA en comparaison aux témoins sains (**Article 4**). En second lieu, nous avons combiné les approches génétiques et génomiques afin d'identifier des polymorphismes associés au niveau d'expression génique d'*ERAP1* (**Article 5**).

### Positionnement du sujet et problématique général

Si les GWAS ont permis l'identification de nombreux facteurs génétiques de susceptibilité aux maladies complexes, la plupart des variants mis en cause exercent une influence modeste sur la prédisposition génétique totale à ces maladies<sup>177</sup>. De plus, dans la majorité des cas, les variants causaux n'ont pas été formellement identifiés et l'apport des polymorphismes associés sur la compréhension de la physiopathologie des maladies étudiées reste globalement faible<sup>465</sup>. Enfin, les études génétiques permettent difficilement la prise en compte de facteurs environnementaux potentiellement impliqués dans la survenue des maladies complexes<sup>466</sup>.

Grâce aux progrès technologiques, l'étude de l'expression des gènes s'est développée et constitue désormais une approche complémentaire à celle de la génétique classique. En effet, tout comme les GWAS, elle permet une approche haut-débit pangénomique et ne nécessite donc pas d'émettre des hypothèses physiopathologiques *a priori*. En outre, elle présente un caractère

plus fonctionnel car elle reflète à la fois le terrain génétique sous-jacent mais aussi les facteurs environnementaux auquel le patient est exposé.

Concernant la SpA, plus d'une dizaine d'études d'expression pangénomique ont été publiées avec une grande hétérogénéité de méthodologie<sup>386-398</sup>. Par conséquent, il n'est pas surprenant que les résultats obtenus soient très différents d'une étude à l'autre. La majorité des études se sont intéressées au profil d'expression d'ARN extraits à partir de sang total ou des PBMCs<sup>386-388,390,391,393,395</sup>. Plus rarement, ce sont des tissus potentiellement impliqués dans la maladie qui ont été étudiés<sup>389,392,394,397,398</sup> ou un type cellulaire précis comme le macrophage<sup>396</sup>. Compte-tenu des résultats antérieurs obtenus chez le rat qui soulignent l'implication potentielle des cellules dendritiques dans la SpA, nous avons étudié le transcriptome de MD-DCs avant et après stimulation par le LPS et nous l'avons comparé à celui de témoins sains (**Article 4**). Nous avons enfin cherché à combiner les études génétiques et transcriptomiques dans une approche de type gène-candidat centrée sur *ERAP1* (**Article 5**).

## Étude du transcriptome des cellules dendritiques dans la spondyloarthrite

### Cellules dendritiques et spondyloarthrite

Le choix d'étudier les DCs dans la SpA a découlé des résultats obtenus dans le modèle animal de SpA du rat transgénique HLA-B27. Dans ce modèle, de nombreuses anomalies fonctionnelles des cellules dendritiques ont été mises en évidence<sup>410-413</sup>. Nous avons donc cherché à savoir si des anomalies similaires à celles observées chez le rat HLA-B27+ étaient présentes chez les patients atteints de SpA.

Contrairement aux études menées chez le rat où les DCs provenant de la rate étaient étudiées *ex vivo*, nous avons choisi d'étudier chez l'homme des cellules dendritiques dérivées des monocytes du sang circulant. Ce choix est justifié par deux raisons. En premier lieu, les DCs constituent une fraction très minoritaire des PBMCs (<1%)<sup>467</sup>. Si nous avions directement sélectionné des DCs à partir du sang circulant, nous n'aurions pas pu obtenir suffisamment de cellules pour réaliser simultanément l'ensemble des études que nous avons prévues (stimulation des lymphocytes T, cytométrie de flux et transcriptome). Un autre avantage à étudier des cellules différenciées *in vitro*, comme les MD-DCs, est qu'elles échappent à l'influence directe de l'environnement systémique inflammatoire ou des traitements médicamenteux<sup>396</sup>.

Nous nous sommes d'abord assurés que les monocytes à partir desquels ont été dérivées les cellules dendritiques présentaient un phénotype comparable entre les patients et les témoins. Ensuite la capacité des MD-DCs de patients atteints de SpA à induire une réponse proliférative de LT

CD4+ hétérologues a été comparée à celle de MD-DCs provenant de témoins. Une diminution significative a été observée avec les MD-DCs de patients laissées au repos ou stimulées pendant six heures par le LPS. Les résultats obtenus suggéraient aussi que la cinétique de réponse des MD-DCs à la stimulation par le LPS était différente. Ainsi, chez les témoins, on observait une augmentation de l'indice de prolifération des LT après six heures de stimulation des MD-DCs par le LPS puis un retour presque à l'état de base après 24 heures de stimulation. Au contraire, chez les patients, l'indice de prolifération des LT variait très peu entre les MD-DCs laissées au repos et celles ayant reçu six heures de stimulation par le LPS, alors qu'il augmentait avec les MD-DCs stimulées durant 24 heures.

Afin de tenter d'expliquer les anomalies fonctionnelles observées, une analyse en cytométrie de flux a été réalisée et n'a montré aucune différence significative de l'expression des marqueurs étudiés à la surface cellulaire des MD-DCs (CD40, HLA-I, HLA-DR, CD80, CD83 et CD86) entre patients et témoins. Nos résultats diffèrent de ceux publiés par Slobodin *et al.* où l'expression du HLA-DR était inférieure à la surface de DCs non stimulées provenant de sujets atteints de SA en comparaison à celles de témoins sains. Cette différence disparaissait toutefois après stimulation<sup>468</sup>.

### Transcriptome des cellules dendritiques

Pour tenter de comprendre l'origine des anomalies fonctionnelles mises en évidence par l'étude des MD-DCs de patients atteints de SpA, nous avons entrepris d'étudier leur profil d'expression génique. En raison des variations de cinétique observées lors des études de prolifération, nous avons opté pour une étude comparative dynamique des MD-DCs de patients et de témoins avant et après stimulation par le LPS. Il s'agit d'un dessin expérimental original car, si certaines études ont analysé les variations de l'expression génique à différents temps<sup>469</sup>, peu d'entre elles ont intégré à la fois un facteur catégoriel (ici cas/témoins) et un facteur temporel<sup>470,471</sup>.

Le premier point à souligner concernant les résultats de cette étude est que l'effet sur la variation du niveau d'expression génique de la stimulation par le LPS était plus important que l'effet lié à la maladie. En effet le clustering hiérarchique non supervisé réalisé à partir des données pangénomiques permettait de distinguer facilement les trois temps de stimulation mais ne permettait pas de séparer les malades des témoins.

En appliquant une correction rigoriste pour tests multiples telle que la méthode de Bonferroni ou le FDR, aucun gène n'était différentiellement exprimé entre patients et témoins. La principale raison permettant d'expliquer ce résultat négatif en était la puissance insuffisante de notre étude due à des effectifs relativement limités (neuf malades et dix témoins). Pour poursuivre l'exploitation de nos résultats, nous avons donc décidé de sélectionner les gènes différentiellement

exprimés sur des critères moins stricts tout en gardant à l'esprit que la validité statistique de nos résultats était imparfaite. En appliquant un filtre de p-value  $<0,01$  et de fold-change  $P \leq -1,5$  ou  $\geq 1,5$ , 81 gènes se sont révélés différentiellement exprimés entre patients et témoins. Comme attendu, la classification hiérarchique non supervisée à partir de cette liste de gènes permettait une bonne séparation entre les patients et les témoins. Une analyse des voies biologiques potentiellement impliquées réalisée à partir de cette liste n'a en revanche pas mis en évidence de voie spécifiquement dérégulée. Notamment, nous n'avons pas retrouvé de signature interféron gamma inverse, comme cela avait été mis en évidence dans les DCs de rats transgéniques B27 ainsi que dans les macrophages de patients atteints de SpA<sup>396,402</sup>.

Parmi ces 81 gènes, nous en avons sélectionnés quatre pour une validation technique par PCR quantitative : *ADAMTS15*, *CITED2*, *F13A1* et *SELL*. Le choix s'est effectué sur la base de la valeur p, du fold-change et de la pertinence biologique. *ADAMTS15*, *F13A1* et *SELL* ont tous été impliqués dans des processus inflammatoires<sup>472-474</sup>. Néanmoins, il existe peu de données concernant leur rôle spécifique dans les DCs.

En ce qui concerne le gène *CITED2*, il exerce un rôle de co-activateur transcriptionnel et peut donc, en fonction du contexte, réguler positivement ou bien négativement l'expression d'un grand nombre de gènes. Nous avons ainsi pu montrer une corrélation inverse entre son niveau d'expression et celui d'*ADAMTS15* dans les MD-DCs stimulées pendant six heures par le LPS. Des corrélations négatives similaires ont été rapportées entre *CITED2* et plusieurs membres de la famille des métalloprotéinases dont fait partie *ADAMTS15*, renforçant la validité de nos résultats<sup>475</sup>.

Considérant son rôle de co-activateur transcriptionnel, nous avons cherché à identifier les gènes dont l'expression était corrélée à *CITED2*. Nous avons cherché à mettre en évidence des différences de co-expression génique entre malades et témoins. Ces analyses ont souligné une forte corrélation de *CITED2* avec plusieurs gènes de la voie Wnt chez les malades alors qu'une telle co-expression n'était pas retrouvée chez les témoins. Outre son rôle connu dans la formation osseuse<sup>476</sup>, il a récemment été montré que l'activation de la voie Wnt dans les DCs modifiait l'équilibre entre LT régulateurs et lymphocytes de la voie Th17<sup>477</sup>.

### Perspectives

Avant d'approfondir l'étude des gènes d'intérêt identifiés dans cette étude transcriptomique, il nous a paru indispensable de confirmer ces premiers résultats. Nous avons donc entrepris une seconde étude utilisant le même type de puces ARN et le même dessin expérimental. Pour ce faire, nous avons obtenu de nouveaux échantillons de MD-DCs provenant de 13 malades et de 11 témoins

indépendants HLA-B27– supplémentaires. Nous avons également complété cette cohorte par le recrutement de témoins sains HLA-B27+ appartenant à la fratrie des malades ainsi que de témoins HLA-B27+ indépendants. Nous espérons ainsi pouvoir déterminer l'influence respective du HLA-B27 et de la maladie sur le profil d'expression génique. Le recrutement d'apparentés devrait en outre permettre de limiter la variabilité d'expression génique inter-individuelle et de gagner en puissance statistique<sup>478</sup>. Enfin, la plupart des paires de germains recrutées appartenant à des familles multiplex analysées dans l'étude de liaison, nous espérons pouvoir identifier des différences d'expression génique correspondant aux régions chromosomiques liées à la SpA et nous aider ainsi de ces variations d'expression pour identifier le(s) gène(s) en cause.

## Du polymorphisme à la fonction : exemple du gène *ERAP1*

### *Du polymorphisme à l'expression génique*

Les GWAS ont permis l'identification d'un grand nombre de polymorphismes associés aux maladies complexes. Toutefois, il est souvent difficile d'établir les conséquences fonctionnelles de ces variants. Parmi les mécanismes en cause, certains polymorphismes non synonymes altèrent la fonction de la protéine codée par le gène à l'intérieur duquel ils sont situés : c'est le cas par exemple des polymorphismes de NOD2 associés aux MICI<sup>479</sup>. Une autre conséquence fonctionnelle possible, en est une modification de l'expression d'un gène par l'action directe ou indirecte d'un SNP sur des éléments régulateurs : on parle alors d'eQTLs<sup>480</sup>. L'investigation systématique des SNPs découverts dans les GWAS a révélé un enrichissement en eQTLs, témoignant de l'importance de ce mécanisme dans la pathogénie des maladies complexes<sup>481</sup>.

Dans la SpA, peu de conséquences fonctionnelles ont été décrites pour les polymorphismes associés à la maladie, en dehors des gènes *IL23R* et *ERAP1*. La plupart des variants d'*ERAP1* associés à la maladie sont non synonymes et il a été notamment suggéré que le SNP rs30187 altérerait la fonction enzymatique d'ERAP1 en modifiant la conformation tri-dimensionnelle de l'enzyme. Nous avons démontré que les haplotypes d'*ERAP1* modulant la susceptibilité à la SpA influençaient le niveau d'expression du transcrit d'ERAP1, suggérant ainsi un second mécanisme fonctionnel possible à l'origine de l'association génétique *ERAP1*/SpA. Il s'agit à notre connaissance de la première démonstration d'un eQTL associé à la SpA.

### *Du gène à la protéine et à la fonction*

Nous avons ensuite montré que les haplotypes d'ERAP1 associés à la SpA influençaient aussi le niveau d'expression protéique dans le même sens que celui observé pour les transcrits, faisant d'ERAP1 un pQTL (locus associé à un trait quantitatif protéique) sous le contrôle de cet haplotype.

Cette vérification était importante, car les variations de l'expression de l'ARNm d'un gène n'impliquent pas forcément des variations similaires de l'expression de la protéine qu'il code, en raison de possibles modifications post-transcriptionnelles. Ainsi dans une étude systématique pangénomique menée par Wu *et al.*, seuls la moitié des eQTLs mis en évidence s'avéraient être également des pQTLs<sup>482</sup>.

La dernière étape de notre étude a cherché à déterminer dans quelle mesure les modifications du niveau d'expression protéique induites par les différents haplotypes d'*ERAP1* pourraient expliquer des différences de niveau d'activité enzymatique mises en évidence dans d'autres études<sup>269,357–360,483</sup>. Nous avons observé que l'activité enzymatique d'*ERAP1* était nettement inférieure chez les porteurs de l'haplotype protecteur, mais que les porteurs de l'haplotype neutre et de l'haplotype à risque ne présentaient pas de différences majeures. Les polymorphismes d'*ERAP1* associés à la SpA semblent donc moduler le niveau d'activité enzymatique d'*ERAP1* en agissant à la fois sur la quantité d'enzyme produite et sur la conformation de celle-ci.

Notre étude permet ainsi de mieux comprendre les conséquences fonctionnelles des polymorphismes d'*ERAP1* associés à la SpA. Il reste néanmoins à expliquer quelles sont les conséquences fonctionnelles de la modification du niveau d'activité enzymatique d'*ERAP1* et en quoi ces conséquences pourraient favoriser la survenue de la SpA. Tenant compte de l'interaction génétique décrite entre *ERAP1* et HLA-B27 – l'association entre *ERAP1* et SpA semble restreinte aux porteurs de l'allèle B27 –, la première hypothèse était que les variations de l'activité enzymatique d'*ERAP1* se répercutent sur l'expression du CMH de classe I à la surface cellulaire. Nous n'avons toutefois pas confirmé cette hypothèse. Une autre hypothèse validée par plusieurs études serait que le niveau d'activité enzymatique d'*ERAP1* puisse influencer qualitativement, et non plus quantitativement, le peptidome présentés par le CMH de classe I et notamment par HLA-B27<sup>359,362,363</sup>. Il reste toutefois à comprendre par quel(s) mécanisme(s) de telles modifications qualitatives pourraient à leur tour influencer sur le développement de la SpA.

### Perspectives

L'approche utilisée dans l'**Article 5** combinant génétique et génomique a permis d'identifier les conséquences fonctionnelles de certains SNPs d'*ERAP1* et ainsi d'avancer dans la compréhension des mécanismes liant *ERAP1* à la SpA. Nous envisageons de poursuivre ce type d'approche à la recherche d'autres eQTLs impliqués dans la SpA. Pour ce faire, l'ensemble des individus inclus dans nos études transcriptomiques seront prochainement génotypés à l'aide de la puce Immunochip dédiée à l'étude des gènes candidats des maladies inflammatoires et auto-immunes. Nous pourrons ainsi mener une étude systématique de corrélation génotype/expression.

## **QUATRIÈME PARTIE : CONCLUSIONS**

## Conclusions

---

Malgré la découverte récente de plus d'une vingtaine de locus de susceptibilité à la maladie, l'architecture génétique globale de la SpA reste encore largement inconnue. En effet, si HLA-B27 semble être un facteur prépondérant, bien que non indispensable à la survenue de la maladie, on ignore combien d'autres facteurs génétiques sont en cause et dans quelle mesure ils interagissent entre eux ou avec des facteurs environnementaux.

Durant ces quatre années de thèse, nous avons cherché à identifier de nouveaux facteurs génétiques de susceptibilité à la SpA. L'approche génétique familiale nous a permis d'ouvrir de nouvelles pistes avec la mise en évidence d'un locus significativement lié à la maladie en 13q13 et d'un SNP du gène *MAPK14* significativement associé à la SpA. Ces résultats nécessitent toutefois d'être approfondis. Le séquençage de la région 13q13 chez des individus appartenant aux familles les plus liées à ce locus est en cours. Nous espérons ainsi pouvoir identifier le(s) variant(s) à l'origine du signal de liaison. Par ailleurs, le séquençage du gène *MAPK14* couplé à des études fonctionnelles serait intéressant afin de mieux comprendre l'implication de ce gène dans la SpA.

Les études menées sur les DCs de patients atteints de SpA en parallèle des études génétiques valident chez le patient une partie des anomalies fonctionnelles observées chez le rat transgénique HLA-B27. L'analyse du transcriptome de ces cellules a en outre permis de mettre en évidence 81 gènes différentiellement exprimés entre patients et témoins. Ces résultats nécessitent d'être confirmés et une étude de réplication est en cours. Cette étude inclura aussi des témoins HLA-B27+. La combinaison des données génétiques et transcriptomiques appliquées à l'étude du gène *ERAP1* apporte un éclairage nouveau à l'association de ce gène à la SpA. En effet, les résultats obtenus suggèrent que l'expression de ce gène est sous le contrôle de polymorphismes associés à la maladie. Les conséquences de ces variations d'expression restent néanmoins à déterminer.

Les études menées durant cette thèse nous ont permis d'accumuler un très grand nombre de données génétiques et d'expression. Outre l'approfondissement des résultats déjà obtenus, nous comptons poursuivre l'étude de ces données en appliquant les stratégies déjà mises en place, afin d'en tirer le meilleur parti. L'exemple d'*ERAP1* illustre bien la puissance d'une étude combinant les données génétiques à celles d'expression. Nous sommes convaincus que, grâce à ce type d'approche intégrative, nous pourrions avancer plus rapidement dans la compréhension des mécanismes en cause dans la SpA, ce qui devrait permettre des progrès diagnostiques et thérapeutiques significatifs pour cette maladie.



## **CINQUIÈME PARTIE : BIBLIOGRAPHIE**

1. Khan, M. A. HLA-B27 and its subtypes in world populations. *Curr Opin Rheumatol* **7**, 263–269 (1995).
2. Saraux, A. *et al.* Prevalence of spondyloarthropathies in France: 2001. *Ann. Rheum. Dis.* **64**, 1431–1435 (2005).
3. Guillemin, F. *et al.* Prevalence of rheumatoid arthritis in France: 2001. *Ann. Rheum. Dis.* **64**, 1427–1430 (2005).
4. Costantino, F. *et al.* Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann. Rheum. Dis.* (2013). doi:10.1136/annrheumdis-2013-204436
5. Lee, W. *et al.* Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. *Ann. Rheum. Dis.* **66**, 633–638 (2007).
6. Feldtkeller, E., Bruckel, J. & Khan, M. A. Scientific contributions of ankylosing spondylitis patient advocacy groups. *Curr Opin Rheumatol* **12**, 239–247 (2000).
7. Stolwijk, C., Boonen, A., van Tubergen, A. & Reveille, J. D. Epidemiology of spondyloarthritis. *Rheum. Dis. Clin. North Am.* **38**, 441–476 (2012).
8. Boyer, G. S., Lanier, A. P. & Templin, D. W. Prevalence rates of spondyloarthropathies, rheumatoid arthritis, and other rheumatic disorders in an Alaskan Inupiat Eskimo population. *J. Rheumatol.* **15**, 678–683 (1988).
9. Alexeeva, L. *et al.* Prevalence of spondyloarthropathies and HLA-B27 in the native population of Chukotka, Russia. *J. Rheumatol.* **21**, 2298–2300 (1994).
10. Braun, J. *et al.* Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. *Arthritis Rheum.* **41**, 58–67 (1998).
11. Farooqi, A. & Gibson, T. Prevalence of the major rheumatic disorders in the adult population of north Pakistan. *Br. J. Rheumatol.* **37**, 491–495 (1998).
12. Hukuda, S. *et al.* Spondyloarthropathies in Japan: nationwide questionnaire survey performed by the Japan Ankylosing Spondylitis Society. *J. Rheumatol.* **28**, 554–559 (2001).
13. Bruges-Armas, J. *et al.* Prevalence of spondyloarthritis in Terceira, Azores: a population based study. *Ann. Rheum. Dis.* **61**, 551–553 (2002).
14. Minh Hoa, T. T. *et al.* Prevalence of the rheumatic diseases in urban Vietnam: a WHO-ILAR COPCORD study. *J. Rheumatol.* **30**, 2252–2256 (2003).
15. Trontzas, P. *et al.* Seronegative spondyloarthropathies in Greece: a population-based study of prevalence, clinical pattern, and management. The ESORDIG study. *Clin. Rheumatol.* **24**, 583–589 (2005).
16. De Angelis, R., Salaffi, F. & Grassi, W. Prevalence of spondyloarthropathies in an Italian population sample: a regional community-based study. *Scand. J. Rheumatol.* **36**, 14–21 (2007).

17. Adomaviciute, D. *et al.* Prevalence survey of rheumatoid arthritis and spondyloarthropathy in Lithuania. *Scand. J. Rheumatol.* **37**, 113–119 (2008).
18. Davatchi, F. *et al.* WHO-ILAR COPCORD Study (Stage 1, Urban Study) in Iran. *J. Rheumatol.* **35**, 1384 (2008).
19. Onen, F. *et al.* Prevalence of ankylosing spondylitis and related spondyloarthritides in an urban area of Izmir, Turkey. *J. Rheumatol.* **35**, 305–309 (2008).
20. Liao, Z. T. *et al.* An epidemiological survey of low back pain and axial spondyloarthritis in a Chinese Han population. *Scand. J. Rheumatol.* **38**, 455–459 (2009).
21. Haglund, E. *et al.* Prevalence of spondyloarthritis and its subtypes in southern Sweden. *Ann Rheum Dis* **70**, 943–948 (2011).
22. Reveille, J. D., Witter, J. P. & Weisman, M. H. Prevalence of axial spondylarthritis in the United States: estimates from a cross-sectional survey. *Arthritis Care Res (Hoboken)* **64**, 905–910 (2012).
23. Van der Linden, S. M., Valkenburg, H. A., de Jongh, B. M. & Cats, A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum.* **27**, 241–249 (1984).
24. Salvadorini, G. *et al.* Ankylosing spondylitis: how diagnostic and therapeutic delay have changed over the last six decades. *Clin. Exp. Rheumatol.* **30**, 561–565 (2012).
25. Sørensen, J., Hetland, M. L. & on behalf of all departments of rheumatology in Denmark. Diagnostic delay in patients with rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis: results from the Danish nationwide DANBIO registry. *Ann. Rheum. Dis.* (2014). doi:10.1136/annrheumdis-2013-204867
26. Dougados, M., Kahan, A. & Revel, M. *La spondylarthrite en 100 questions*. (Paris : AP-HP Assistance Publique-Hopitaux de Paris, 2005).
27. Appel, H. & Sieper, J. Spondyloarthritis at the crossroads of imaging, pathology, and structural damage in the era of biologics. *Curr Rheumatol Rep* **10**, 356–363 (2008).
28. Eshed, I. *et al.* MRI of enthesitis of the appendicular skeleton in spondyloarthritis. *Annals of the Rheumatic Diseases* **66**, 1553–1559 (2007).
29. Rudwaleit, M. *et al.* The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. *Arthritis Rheum.* **60**, 717–727 (2009).
30. Dougados, M. *et al.* The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. *Joint Bone Spine* **78**, 598–603 (2011).
31. Maksymowych, W. P. *et al.* Development and validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index. *Ann. Rheum. Dis.* **68**, 948–953 (2009).

32. Canouï-Poitrine, F. *et al.* Prevalence and factors associated with uveitis in spondylarthritis patients in France: results from an observational survey. *Arthritis Care Res (Hoboken)* **64**, 919–924 (2012).
33. Zeboulon, N., Dougados, M. & Gossec, L. Prevalence and characteristics of uveitis in the spondyloarthropathies: a systematic literature review. *Ann. Rheum. Dis.* **67**, 955–959 (2008).
34. Gelfand, J. M. *et al.* Prevalence and treatment of psoriasis in the United Kingdom: a population-based study. *Arch Dermatol* **141**, 1537–1541 (2005).
35. Griffiths, C. E. M. & Barker, J. N. W. N. Pathogenesis and clinical features of psoriasis. *Lancet* **370**, 263–271 (2007).
36. Wright, V., Roberts, M. C. & Hill, A. G. Dermatological manifestations in psoriatic arthritis: a follow-up study. *Acta Derm. Venereol.* **59**, 235–240 (1979).
37. Gladman, D. D. Psoriatic arthritis. *Rheum. Dis. Clin. North Am.* **24**, 829–844, x (1998).
38. Mielants, H. *et al.* The evolution of spondyloarthropathies in relation to gut histology. II. Histological aspects. *J. Rheumatol.* **22**, 2273–2278 (1995).
39. El Maghraoui, A. Extra-articular manifestations of ankylosing spondylitis: prevalence, characteristics and therapeutic implications. *Eur. J. Intern. Med.* **22**, 554–560 (2011).
40. Olhagen, B. Urogenital syndromes and spondarthritis. *Br. J. Rheumatol.* **22**, 33–40 (1983).
41. Rudwaleit, M., van der Heijde, D., Khan, M. A., Braun, J. & Sieper, J. How to diagnose axial spondyloarthritis early. *Ann. Rheum. Dis.* **63**, 535–543 (2004).
42. Gran, J. T. & Husby, G. HLA-B27 and spondyloarthropathy: value for early diagnosis? *J. Med. Genet.* **32**, 497–501 (1995).
43. Dougados, M. *et al.* Clinical relevance of C-reactive protein in axial involvement of ankylosing spondylitis. *J. Rheumatol.* **26**, 971–974 (1999).
44. Lukas, C. *et al.* Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann. Rheum. Dis.* **68**, 18–24 (2009).
45. Van der Linden, S., Valkenburg, H. A. & Cats, A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum.* **27**, 361–368 (1984).
46. Sieper, J. *et al.* The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann. Rheum. Dis.* **68 Suppl 2**, ii1–44 (2009).
47. Devauchelle-Pensec, V. *et al.* Computed tomography scanning facilitates the diagnosis of sacroiliitis in patients with suspected spondylarthritis: results of a prospective multicenter French cohort study. *Arthritis Rheum.* **64**, 1412–1419 (2012).
48. Rudwaleit, M., Khan, M. A. & Sieper, J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis Rheum.* **52**, 1000–1008 (2005).

49. Rudwaleit, M. *et al.* The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann. Rheum. Dis.* **68**, 777–783 (2009).
50. Rudwaleit, M. *et al.* Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann. Rheum. Dis.* **68**, 1520–1527 (2009).
51. Hermann, K.-G. A. *et al.* Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann. Rheum. Dis.* **71**, 1278–1288 (2012).
52. D’Agostino, M.-A. *et al.* Assessment of peripheral enthesitis in the spondylarthropathies by ultrasonography combined with power Doppler: a cross-sectional study. *Arthritis Rheum.* **48**, 523–533 (2003).
53. Lehtinen, A., Taavitsainen, M. & Leirisalo-Repo, M. Sonographic analysis of enthesopathy in the lower extremities of patients with spondylarthropathy. *Clin. Exp. Rheumatol.* **12**, 143–148 (1994).
54. Balint, P. V., Kane, D., Wilson, H., McInnes, I. B. & Sturrock, R. D. Ultrasonography of enthesal insertions in the lower limb in spondylarthropathy. *Ann. Rheum. Dis.* **61**, 905–910 (2002).
55. Braun, J. *et al.* 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann. Rheum. Dis.* **70**, 896–904 (2011).
56. Zochling, J. *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann. Rheum. Dis.* **65**, 442–452 (2006).
57. Poddubnyy, D. Axial spondyloarthritis: is there a treatment of choice? *Ther Adv Musculoskelet Dis* **5**, 45–54 (2013).
58. Dougados, M. *et al.* ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann. Rheum. Dis.* **70**, 249–251 (2011).
59. Poddubnyy, D. *et al.* Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. *Ann. Rheum. Dis.* **71**, 1616–1622 (2012).
60. Haibel, H. *et al.* Efficacy of oral prednisolone in active ankylosing spondylitis: results of a double-blind, randomised, placebo-controlled short-term trial. *Ann. Rheum. Dis.* **73**, 243–246 (2014).
61. Chen, J., Veras, M. M. S., Liu, C. & Lin, J. Methotrexate for ankylosing spondylitis. *Cochrane Database Syst Rev* **2**, CD004524 (2013).
62. Van Denderen, J. C. *et al.* Double blind, randomised, placebo controlled study of leflunomide in the treatment of active ankylosing spondylitis. *Ann. Rheum. Dis.* **64**, 1761–1764 (2005).

63. Chen, J. & Liu, C. Sulfasalazine for ankylosing spondylitis. *Cochrane Database Syst Rev* CD004800 (2005). doi:10.1002/14651858.CD004800.pub2
64. Braun, J. *et al.* Use of immunohistologic and in situ hybridization techniques in the examination of sacroiliac joint biopsy specimens from patients with ankylosing spondylitis. *Arthritis Rheum.* **38**, 499–505 (1995).
65. Present, D. H. *et al.* Infliximab for the treatment of fistulas in patients with Crohn's disease. *N. Engl. J. Med.* **340**, 1398–1405 (1999).
66. Brandt, J. *et al.* Successful treatment of active ankylosing spondylitis with the anti-tumor necrosis factor alpha monoclonal antibody infliximab. *Arthritis Rheum.* **43**, 1346–1352 (2000).
67. Van den Bosch, F. *et al.* Effects of a loading dose regimen of three infusions of chimeric monoclonal antibody to tumour necrosis factor alpha (infliximab) in spondyloarthritis: an open pilot study. *Ann. Rheum. Dis.* **59**, 428–433 (2000).
68. Van der Heijde, D. *et al.* Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). *Arthritis Rheum.* **52**, 582–591 (2005).
69. Van der Heijde, D. *et al.* Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* **54**, 2136–2146 (2006).
70. Davis, J. C. *et al.* Sustained durability and tolerability of etanercept in ankylosing spondylitis for 96 weeks. *Ann. Rheum. Dis.* **64**, 1557–1562 (2005).
71. Inman, R. D. *et al.* Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. *Arthritis Rheum.* **58**, 3402–3412 (2008).
72. Landewé, R. *et al.* Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. *Ann. Rheum. Dis.* **73**, 39–47 (2014).
73. Van der Heijde, D. *et al.* Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. *Arthritis Rheum.* **58**, 1324–1331 (2008).
74. Van der Heijde, D. *et al.* Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. *Arthritis Rheum.* **58**, 3063–3070 (2008).
75. Van der Heijde, D. *et al.* Assessment of radiographic progression in the spines of patients with ankylosing spondylitis treated with adalimumab for up to 2 years. *Arthritis Res. Ther.* **11**, R127 (2009).
76. Baraliakos, X., Haibel, H., Listing, J., Sieper, J. & Braun, J. Continuous long-term anti-TNF therapy does not lead to an increase in the rate of new bone formation over 8 years in patients with ankylosing spondylitis. *Ann. Rheum. Dis.* (2013). doi:10.1136/annrheumdis-2012-202698

77. Haroon, N. *et al.* Anti-TNF Therapy Slows Radiographic Progression of Ankylosing Spondylitis. *Arthritis Rheum.* **64**, 339 (2012).
78. Haibel, H., Rudwaleit, M., Listing, J. & Sieper, J. Open label trial of anakinra in active ankylosing spondylitis over 24 weeks. *Ann. Rheum. Dis.* **64**, 296–298 (2005).
79. Song, I.-H. *et al.* Treatment of active ankylosing spondylitis with abatacept: an open-label, 24-week pilot study. *Ann. Rheum. Dis.* **70**, 1108–1110 (2011).
80. Song, I.-H. *et al.* Different response to rituximab in tumor necrosis factor blocker-naïve patients with active ankylosing spondylitis and in patients in whom tumor necrosis factor blockers have failed: a twenty-four-week clinical trial. *Arthritis Rheum.* **62**, 1290–1297 (2010).
81. Sieper, J., Inman, R., Badalamenti, S., Radin, I. & Braun, J. Sarilumab for the treatment of ankylosing spondylitis: results of a phase 2, randomized, double-blind, placebo-controlled, international study (ALIGN). *Ann. Rheum. Dis.* (2014).
82. Sieper, J., Porter-Brown, B., Thompson, L., Harari, O. & Dougados, M. Assessment of short-term symptomatic efficacy of tocilizumab in ankylosing spondylitis: results of randomised, placebo-controlled trials. *Ann. Rheum. Dis.* (2013). doi:10.1136/annrheumdis-2013-203559
83. Baeten, D. *et al.* Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. *Lancet* (2013). doi:10.1016/S0140-6736(13)61134-4
84. Baeten, D. *et al.* The Anti-IL17A monoclonal antibody secukinumab (AIN457) showed good safety and efficacy in the treatment of active ankylosing spondylitis. *Arthritis Rheum.* **62**, 3840 (2010).
85. Gottlieb, A. *et al.* Ustekinumab, a human interleukin 12/23 monoclonal antibody, for psoriatic arthritis: randomised, double-blind, placebo-controlled, crossover trial. *Lancet* **373**, 633–640 (2009).
86. Pathan, E. *et al.* Efficacy and safety of apremilast, an oral phosphodiesterase 4 inhibitor, in ankylosing spondylitis. *Ann. Rheum. Dis.* (2012). doi:10.1136/annrheumdis-2012-201915
87. Poddubnyy, D., Hermann, K.-G. A., Callhoff, J., Listing, J. & Sieper, J. Ustekinumab for the treatment of patients with active ankylosing spondylitis: results of a 28-week, prospective, open-label, proof-of-concept study (TOPAS). *Ann. Rheum. Dis.* (2014). doi:10.1136/annrheumdis-2013-204248
88. Dagfinrud, H., Kvien, T. K. & Hagen, K. B. Physiotherapy interventions for ankylosing spondylitis. *Cochrane Database Syst Rev* CD002822 (2008). doi:10.1002/14651858.CD002822.pub3
89. Wilson, F. C. *et al.* Time trends in epidemiology and characteristics of psoriatic arthritis over 3 decades: a population-based study. *J. Rheumatol.* **36**, 361–367 (2009).
90. Cantini, F. *et al.* Psoriatic arthritis: a systematic review. *Int J Rheum Dis* **13**, 300–317 (2010).
91. Aho, K., Ahvonen, P., Lassus, A., Sievers, K. & Tilikainen, A. HL-A antigen 27 and reactive arthritis. *Lancet* **2**, 157 (1973).

92. Amor, B. Reiter's syndrome. Diagnosis and clinical features. *Rheum. Dis. Clin. North Am.* **24**, 677–695, vii (1998).
93. Shivashankar, R. *et al.* Incidence of spondyloarthropathy in patients with Crohn's disease: a population-based study. *J. Rheumatol.* **39**, 2148–2152 (2012).
94. Shivashankar, R. *et al.* Incidence of Spondyloarthropathy in patients with ulcerative colitis: a population-based study. *J. Rheumatol.* **40**, 1153–1157 (2013).
95. Mau, W. *et al.* Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup. *J. Rheumatol.* **15**, 1109–1114 (1988).
96. Copeman. in *Textbook of rheumatic diseases* 12 (Livingstone, 1967).
97. Moll, J. M., Haslock, I., Macrae, I. F. & Wright, V. Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behcet's syndrome. *Medicine (Baltimore)* **53**, 343–364 (1974).
98. Caffrey, M. F. & James, D. C. Human lymphocyte antigen association in ankylosing spondylitis. *Nature* **242**, 121 (1973).
99. Brewerton, D. A. *et al.* Ankylosing spondylitis and HL-A 27. *Lancet* **1**, 904–907 (1973).
100. Reveille, J. D. The genetic basis of spondyloarthritis. *Ann. Rheum. Dis.* **70** Suppl 1, i44–50 (2011).
101. Amor, B., Dougados, M. & Mijiyawa, M. [Criteria of the classification of spondylarthropathies]. *Rev Rhum Mal Osteoartic* **57**, 85–89 (1990).
102. Dougados, M. *et al.* The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. *Arthritis Rheum.* **34**, 1218–1227 (1991).
103. Amor, B. *et al.* [Evaluation of the Amor criteria for spondylarthropathies and European Spondylarthropathy Study Group (ESSG). A cross-sectional analysis of 2,228 patients]. *Ann Med Interne (Paris)* **142**, 85–89 (1991).
104. Rudwaleit, M. *et al.* The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann. Rheum. Dis.* **70**, 25–31 (2011).
105. Burton, P. R., Tobin, M. D. & Hopper, J. L. Key concepts in genetic epidemiology. *Lancet* **366**, 941–951 (2005).
106. Litt, M. & Luty, J. A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**, 397 (1989).
107. Weber, J. L. & May, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**, 388 (1989).
108. Weissenbach, J. *et al.* A second-generation linkage map of the human genome. *Nature* **359**, 794–801 (1992).



109. Weissenbach, J. Microsatellite polymorphisms and the genetic linkage map of the human genome. *Curr. Opin. Genet. Dev.* **3**, 414–417 (1993).
110. Abecasis, G. R. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
111. Chorley, B. N. *et al.* Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: Current and developing technologies. *Mutat. Res.* **659**, 147–157 (2008).
112. Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human genome. *Nat Rev Genet* **7**, 85–97 (2006).
113. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
114. Duan, J., Zhang, J.-G., Deng, H.-W. & Wang, Y.-P. Comparative Studies of Copy Number Variation Detection Methods for Next-Generation Sequencing Technologies. *PLoS ONE* **8**, e59128 (2013).
115. Beckmann, J. S., Estivill, X. & Antonarakis, S. E. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet* **8**, 639–646 (2007).
116. Dawn Teare, M. & Barrett, J. H. Genetic linkage studies. *Lancet* **366**, 1036–1044 (2005).
117. HALDANE, J. B. S. & SMITH, C. A. B. A new estimate of the linkage between the genes for colourblindness and haemophilia in man. *Ann Eugen* **14**, 10–31 (1947).
118. Hutchinson, J. B. The Application of the ‘Method of Maximum Likelihood’ to the Estimation of Linkage. *Genetics* **14**, 519–537 (1929).
119. MORTON, N. E. Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* **7**, 277–318 (1955).
120. Elston, R. C. & Stewart, J. A general model for the genetic analysis of pedigree data. *Hum. Hered.* **21**, 523–542 (1971).
121. Ott, J. Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. *Am. J. Hum. Genet.* **26**, 588–597 (1974).
122. Lander, E. S. & Green, P. Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2363–2367 (1987).
123. Sobel, E. & Lange, K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am. J. Hum. Genet.* **58**, 1323–1337 (1996).
124. Lathrop, G. M. & Lalouel, J. M. Easy calculations of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* **36**, 460–465 (1984).
125. Kruglyak, L., Daly, M. J., Reeve-Daly, M. P. & Lander, E. S. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.* **58**, 1347–1363 (1996).
126. Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30**, 97–101 (2002).

127. Morton, N. E., MacLean, C. J., Lew, R. & Yee, S. Multipoint linkage analysis. *Am. J. Hum. Genet.* **38**, 868 (1986).
128. Abreu, P. C., Hodge, S. E. & Greenberg, D. A. Quantification of type I error probabilities for heterogeneity LOD scores. *Genet. Epidemiol.* **22**, 156–169 (2002).
129. SMITH, C. A. TESTING FOR HETEROGENEITY OF RECOMBINATION FRACTION VALUES IN HUMAN GENETICS. *Ann. Hum. Genet.* **27**, 175–182 (1963).
130. Hodge, S. E., Anderson, C. E., Neiswanger, K., Sparkes, R. S. & Rimoïn, D. L. The search for heterogeneity in insulin-dependent diabetes mellitus (IDDM): linkage studies, two-locus models, and genetic heterogeneity. *Am. J. Hum. Genet.* **35**, 1139–1155 (1983).
131. Ott, J. Linkage analysis and family classification under heterogeneity. *Ann. Hum. Genet.* **47**, 311–320 (1983).
132. Conneally, P. M. *et al.* Report of the Committee on Methods of Linkage Analysis and Reporting. *Cytogenet. Cell Genet.* **40**, 356–359 (1985).
133. Lander, E. & Kruglyak, L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241–247 (1995).
134. Clerget-Darpoux, F., Bonaïti-Pellié, C. & Hochez, J. Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* **42**, 393–399 (1986).
135. Greenberg, D. A. Inferring mode of inheritance by comparison of lod scores. *Am. J. Med. Genet.* **34**, 480–486 (1989).
136. Hodge, S. E. & Elston, R. C. Lods, wrods, and mods: The interpretation of lod scores calculated under different models. *Genet. Epidemiol.* **11**, 329–342 (1994).
137. Hodge, S. E., Abreu, P. C. & Greenberg, D. A. Magnitude of type I error when single-locus linkage analysis is maximized over models: a simulation study. *Am J Hum Genet* **60**, 217–227 (1997).
138. Greenberg, D. A., Abreu, P. & Hodge, S. E. The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet* **63**, 870–879 (1998).
139. Penrose, L. S. The Detection of Autosomal Linkage in Data Which Consist of Pairs of Brothers and Sisters of Unspecified Parentage. *Annals of Eugenics* **6**, 133–138 (1935).
140. Blackwelder, W. C., Elston, R. C. & Rao, D. C. A comparison of sib-pair linkage tests for disease susceptibility loci. *Genet. Epidemiol.* **2**, 85–97 (1985).
141. Thomson, G. Determining the mode of inheritance of RFLP-associated diseases using the affected sib-pair method. *Am J Hum Genet* **39**, 207 (1986).
142. Lange, K. A test statistic for the affected-sib-set method. *Ann. Hum. Genet.* **50**, 283–290 (1986).
143. Weeks, D. E. & Lange, K. The affected-pedigree-member method of linkage analysis. *Am J Hum Genet* **42**, 315 (1988).

144. Babron, M.-C., Martinez, M., Bonaïti-Pellié, C. & Clerget-Darpoux, F. Linkage detection by the Affected-Pedigree-Member method: What is really tested? *Genet. Epidemiol.* **10**, 389–394 (1993).
145. Risch, N. Genetics of IDDM: evidence for complex inheritance with HLA. *Genet. Epidemiol.* **6**, 143–148 (1989).
146. Risch, N. Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. *Am J Hum Genet* **46**, 242 (1990).
147. Whittemore, A. S. & Halpern, J. A class of tests for linkage using affected pedigree members. *Biometrics* **50**, 118–127 (1994).
148. Whittemore, A. S. & Halpern, J. Probability of gene identity by descent: computation and applications. *Biometrics* **50**, 109–117 (1994).
149. Kong, A. & Cox, N. J. Allele-sharing models: LOD scores and accurate linkage tests. *Am. J. Hum. Genet.* **61**, 1179–1188 (1997).
150. Gudbjartsson, D. F., Jonasson, K., Frigge, M. L. & Kong, A. Allegro, a new computer program for multipoint linkage analysis. *Nat. Genet.* **25**, 12–13 (2000).
151. Witte, J. S., Elston, R. C. & Schork, N. J. Genetic dissection of complex traits. *Nat. Genet.* **12**, 355–356; author reply 357–358 (1996).
152. Curtis, D. Genetic dissection of complex traits. *Nat. Genet.* **12**, 356–358 (1996).
153. Sawcer, S. *et al.* Empirical genomewide significance levels established by whole genome simulations. *Genet. Epidemiol.* **14**, 223–229 (1997).
154. Lin, D. Y. & Zou, F. Assessing genomewide statistical significance in linkage studies. *Genet. Epidemiol.* **27**, 202–214 (2004).
155. Wall, J. D. & Pritchard, J. K. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* **4**, 587–597 (2003).
156. Hirschhorn, J. N. & Daly, M. J. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* **6**, 95–108 (2005).
157. Lewontin, R. C. THE INTERACTION OF SELECTION AND LINKAGE. I. GENERAL CONSIDERATIONS; HETEROTIC MODELS. *Genetics* **49**, 49–67 (1964).
158. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
159. Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. & Lander, E. S. High-resolution haplotype structure in the human genome. *Nat. Genet.* **29**, 229–232 (2001).
160. Jeffreys, A. J., Kauppi, L. & Neumann, R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat. Genet.* **29**, 217–222 (2001).

161. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**, 321–324 (2005).
162. Johnson, G. C. L. *et al.* Haplotype tagging for the identification of common disease genes. *Nat. Genet.* **29**, 233–237 (2001).
163. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
164. Kruglyak, L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22**, 139–144 (1999).
165. Freidlin, B., Zheng, G., Li, Z. & Gastwirth, J. L. Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum. Hered.* **53**, 146–152 (2002).
166. Schaid, D. J., Rowland, C. M., Tines, D. E., Jacobson, R. M. & Poland, G. A. Score Tests for Association between Traits and Haplotypes when Linkage Phase Is Ambiguous. *Am. J. Hum. Genet.* **70**, 425 (2002).
167. Balding, D. J. A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* **7**, 781–791 (2006).
168. Spielman, R. S., McGinnis, R. E. & Ewens, W. J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* **52**, 506–516 (1993).
169. Lake, S. L., Blacker, D. & Laird, N. M. Family-Based Tests of Association in the Presence of Linkage. *The American Journal of Human Genetics* **67**, 1515–1525 (2000).
170. Laird, N. M. & Lange, C. in *Advances in Genetics* (ed. D. C. Rao and C. Charles Gu) **Volume 60**, 219–252 (Academic Press, 2008).
171. Jorgensen, T. J. *et al.* Hypothesis-Driven Candidate Gene Association Studies: Practical Design and Analytical Considerations. *Am. J. Epidemiol.* **170**, 986–993 (2009).
172. Tabor, H. K., Risch, N. J. & Myers, R. M. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* **3**, 391–397 (2002).
173. Bumgarner, R. Overview of DNA microarrays: types, applications, and their future. *Curr Protoc Mol Biol* **Chapter 22**, Unit 22.1. (2013).
174. Klein, R. J. *et al.* Complement Factor H Polymorphism in Age-Related Macular Degeneration. *Science* **308**, 385–389 (2005).
175. Hindorff, L. *et al.* A Catalog of Published Genome-Wide Association Studies. at [www.genome.gov/gwastudies](http://www.genome.gov/gwastudies)
176. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517 (1996).
177. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).

178. Zeggini, E. *et al.* An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets. *Nat. Genet.* **37**, 1320–1322 (2005).
179. Iyengar, S. K. & Elston, R. C. The genetic basis of complex traits: rare variants or ‘common gene, common disease’? *Methods Mol. Biol.* **376**, 71–84 (2007).
180. Rivas, M. A. *et al.* Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat. Genet.* **43**, 1066–1073 (2011).
181. Kiezun, A. *et al.* Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44**, 623–630 (2012).
182. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
183. Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. GRR: graphical representation of relationship errors. *Bioinformatics* **17**, 742–743 (2001).
184. Turner, S. *et al.* in *Current Protocols in Human Genetics* (John Wiley & Sons, Inc., 2001). at <<http://onlinelibrary.wiley.com/doi/10.1002/0471142905.hg0119s68/abstract>>
185. Hardy, G. H. MENDELIAN PROPORTIONS IN A MIXED POPULATION. *Science* **28**, 49–50 (1908).
186. Weinberg, W. Über den Nachweis der Vererbung beim Menschen. . *Wu' rttemb.* 64: 369–382. *Jahresh. Ver. Vaterl. Naturkd* **64**, 369–382 (1908).
187. Hosking, L. *et al.* Detection of genotyping errors by Hardy–Weinberg equilibrium testing. *Eur. J. Hum. Genet.* **12**, 395–399 (2004).
188. Tiet, L. & Cambien, F. Departure from Hardy–Weinberg equilibrium should be systematically tested in studies of association between genetic markers and disease. *Circulation* **92**, 3364–3365 (1995).
189. Wittke-Thompson, J. K., Pluzhnikov, A. & Cox, N. J. Rational Inferences about Departures from Hardy–Weinberg Equilibrium. *Am J Hum Genet* **76**, 967–986 (2005).
190. Freedman, M. L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388–393 (2004).
191. Cardon, L. R. & Palmer, L. J. Population stratification and spurious allelic association. *The Lancet* **361**, 598–604 (2003).
192. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
193. Pritchard, J. K., Stephens, M., Rosenberg, N. A. & Donnelly, P. Association mapping in structured populations. *Am. J. Hum. Genet.* **67**, 170–181 (2000).
194. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
195. Rice, T. K., Schork, N. J. & Rao, D. C. Methods for handling multiple testing. *Adv. Genet.* **60**, 293–308 (2008).

196. Bonferroni, C. in *Studi in Onore del Professore Salvatore Ortu Carboni* 13–60 (1935).
197. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilit . *Pubblicazioni del R Istituto Superiore de Scienze Economiche e Commerciali de Firenze* **8**, 3–62 (1936).
198. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc. B.* **57**, 289–300 (1995).
199. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29**, 1165–1188 (2001).
200. Bertone, P. *et al.* Global Identification of Human Transcribed Sequences with Genome Tiling Arrays. *Science* **306**, 2242–2246 (2004).
201. Shendure, J. The beginning of the end for microarrays? *Nature Methods* **5**, 585–587 (2008).
202. Okoniewski, M. J. & Miller, C. J. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics* **7**, 276 (2006).
203. Draghici, S., Khatri, P., Eklund, A. C. & Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends in Genetics* **22**, 101–109 (2006).
204. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57–63 (2009).
205. Sun, W. & Hu, Y. eQTL Mapping Using RNA-seq Data. *Stat Biosci* **5**, 198–219 (2013).
206. Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* **12**, 87–98 (2011).
207. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
208. Allison, D. B., Cui, X., Page, G. P. & Sabripour, M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* **7**, 55–65 (2006).
209. Cheung, V. G. & Spielman, R. S. Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nat Rev Genet* **10**, 595–604 (2009).
210. Dimas, A. S. *et al.* Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325**, 1246–1250 (2009).
211. Powell, J. E. *et al.* Genetic control of gene expression in whole blood and lymphoblastoid cell lines is largely independent. *Genome Res.* **22**, 456–466 (2012).
212. Hernandez, D. G. *et al.* Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. *Neurobiol. Dis.* **47**, 20–28 (2012).
213. Feingold, E. Methods for Linkage Analysis of Quantitative Trait Loci in Humans. *Theoretical Population Biology* **60**, 167–180 (2001).

214. Boerwinkle, E., Chakraborty, R. & Sing, C. F. The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann. Hum. Genet.* **50**, 181–194 (1986).
215. Ewens, W. J., Li, M. & Spielman, R. S. A review of family-based tests for linkage disequilibrium between a quantitative trait and a genetic marker. *PLoS Genet.* **4**, e1000180 (2008).
216. Pastinen, T. *et al.* Mapping common regulatory variants to human haplotypes. *Hum. Mol. Genet.* **14**, 3963–3971 (2005).
217. TEGNER, W. & LLOYD, K. Ankylosing spondylitis in all three children of one family. *Lancet* **2**, 196 (1949).
218. JACOBS, J. H. & ROSE, F. C. The familial occurrence of ankylosing spondylitis. *Br Med J* **2**, 1139–1140 (1954).
219. Blécourt, J. J. de, Polman, A., Blécourt-Meindersma, T. de, Erlee, T. J. D. & Drion, E. F. Hereditary Factors in Rheumatoid Arthritis and Ankylosing Spondylitis. *Ann Rheum Dis* **20**, 215–223 (1961).
220. Emery, A. E. & Lawrence, J. S. Genetics of ankylosing spondylitis. *J Med Genet* **4**, 239–244 (1967).
221. Risch, N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am. J. Hum. Genet.* **46**, 222–228 (1990).
222. Brown, M. A., Laval, S. H., Brophy, S. & Calin, A. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann. Rheum. Dis.* **59**, 883–886 (2000).
223. Risch, N. The Genetic Epidemiology of Cancer Interpreting Family and Twin Studies and Their Implications for Molecular Genetic Approaches. *Cancer Epidemiol Biomarkers Prev* **10**, 733–741 (2001).
224. Järvinen, P. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis Rheum.* **38**, 381–383 (1995).
225. Brown, M. A. *et al.* Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum.* **40**, 1823–1828 (1997).
226. Pedersen, O. B. *et al.* Ankylosing spondylitis in Danish and Norwegian twins: occurrence and the relative importance of genetic vs. environmental effectors in disease causation. *Scand. J. Rheumatol.* **37**, 120–126 (2008).
227. Reveille, J. D. Genetics of spondyloarthritis--beyond the MHC. *Nat Rev Rheumatol* **8**, 296–304 (2012).
228. Said-Nahal, R. *et al.* The familial form of spondylarthropathy: a clinical study of 115 multiplex families. Groupe Français d'Etude Génétique des Spondylarthropathies. *Arthritis Rheum.* **43**, 1356–1365 (2000).
229. Said-Nahal, R. *et al.* Phenotypic diversity is not determined by independent genetic factors in familial spondylarthropathy. *Arthritis Rheum.* **45**, 478–484 (2001).
230. Porcher, R. *et al.* Two major spondylarthropathy phenotypes are distinguished by pattern analysis in multiplex families. *Arthritis Rheum.* **53**, 263–271 (2005).

231. Dernis, E. *et al.* Recurrence of spondylarthropathy among first-degree relatives of patients: a systematic cross-sectional study. *Ann. Rheum. Dis.* **68**, 502–507 (2009).
232. Breban, M., Hacquard-Bouder, C. & Falgarone, G. Animal models of HLA-B27-associated diseases. *Curr. Mol. Med.* **4**, 31–40 (2004).
233. Taurog, J. D. *et al.* The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *J. Exp. Med.* **180**, 2359–2364 (1994).
234. Rath, H. C. *et al.* Normal luminal bacteria, especially *Bacteroides* species, mediate chronic colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats. *J. Clin. Invest.* **98**, 945–953 (1996).
235. Chung, H. Y., Machado, P., van der Heijde, D., D’Agostino, M.-A. & Dougados, M. Smokers in early axial spondyloarthritis have earlier disease onset, more disease activity, inflammation and damage, and poorer function and health-related quality of life: results from the DESIR cohort. *Ann. Rheum. Dis.* **71**, 809–816 (2012).
236. Poddubnyy, D. *et al.* Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. *Arthritis Rheum.* **64**, 1388–1398 (2012).
237. Zeboulon-Ktorza, N. *et al.* Influence of environmental factors on disease activity in spondyloarthritis: a prospective cohort study. *J. Rheumatol.* **40**, 469–475 (2013).
238. Fernando, M. M. A. *et al.* Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* **4**, e1000024 (2008).
239. Díaz-Peña, R., López-Vázquez, A. & López-Larrea, C. Old and new HLA associations with ankylosing spondylitis. *Tissue Antigens* **80**, 205–213 (2012).
240. Reveille, J. D. Major histocompatibility genes and ankylosing spondylitis. *Best Pract Res Clin Rheumatol* **20**, 601–609 (2006).
241. López-Larrea, C. *et al.* HLA-B27 subtypes in Asian patients with ankylosing spondylitis. Evidence for new associations. *Tissue Antigens* **45**, 169–176 (1995).
242. Paladini, F. *et al.* Distribution of HLA-B27 subtypes in Sardinia and continental Italy and their association with spondylarthropathies. *Arthritis Rheum.* **52**, 3319–3321 (2005).
243. Neefjes, J., Jongsma, M. L. M., Paul, P. & Bakke, O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 (2011).
244. Anfossi, N. *et al.* Human NK cell education by inhibitory receptors for MHC class I. *Immunity* **25**, 331–342 (2006).
245. Van Gaalen, F. A. *et al.* Epistasis between two HLA antigens defines a subset of individuals at a very high risk for ankylosing spondylitis. *Ann. Rheum. Dis.* **72**, 974–978 (2013).



246. Díaz-Peña, R. *et al.* Fine mapping of a major histocompatibility complex in ankylosing spondylitis: association of the HLA-DPA1 and HLA-DPB1 regions. *Arthritis Rheum.* **63**, 3305–3312 (2011).
247. Díaz-Peña, R. *et al.* Genetic study confirms association of HLA-DPA1(\*)01:03 subtype with ankylosing spondylitis in HLA-B27-positive populations. *Hum. Immunol.* **74**, 764–767 (2013).
248. Zhou, X. *et al.* MICA, a gene contributing strong susceptibility to ankylosing spondylitis. *Ann. Rheum. Dis.* (2013). doi:10.1136/annrheumdis-2013-203352
249. Benjamin, R. & Parham, P. Guilt by association: HLA-B27 and ankylosing spondylitis. *Immunol. Today* **11**, 137–142 (1990).
250. Schwimmbeck, P. L., Yu, D. T. & Oldstone, M. B. Autoantibodies to HLA B27 in the sera of HLA B27 patients with ankylosing spondylitis and Reiter's syndrome. Molecular mimicry with *Klebsiella pneumoniae* as potential mechanism of autoimmune disease. *J. Exp. Med.* **166**, 173–181 (1987).
251. Colbert, R. A. The immunobiology of HLA-B27: variations on a theme. *Curr. Mol. Med.* **4**, 21–30 (2004).
252. Chan, A. T., Kollnberger, S. D., Wedderburn, L. R. & Bowness, P. Expansion and enhanced survival of natural killer cells expressing the killer immunoglobulin-like receptor KIR3DL2 in spondylarthritis. *Arthritis Rheum.* **52**, 3586–3595 (2005).
253. Bowness, P. *et al.* Th17 cells expressing KIR3DL2+ and responsive to HLA-B27 homodimers are increased in ankylosing spondylitis. *J. Immunol.* **186**, 2672–2680 (2011).
254. Hamersma, J. *et al.* Is disease severity in ankylosing spondylitis genetically determined? *Arthritis Rheum.* **44**, 1396–1400 (2001).
255. Brophy, S. *et al.* Concordance of disease severity among family members with ankylosing spondylitis? *J. Rheumatol.* **31**, 1775–1778 (2004).
256. International Genetics of Ankylosing Spondylitis Consortium (IGAS) *et al.* Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat. Genet.* **45**, 730–738 (2013).
257. Brown, M. A. *et al.* A genome-wide screen for susceptibility loci in ankylosing spondylitis. *Arthritis Rheum.* **41**, 588–595 (1998).
258. Laval, S. H. *et al.* Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. *Am. J. Hum. Genet.* **68**, 918–926 (2001).
259. Zhang, G. *et al.* Genetic studies in familial ankylosing spondylitis susceptibility. *Arthritis Rheum.* **50**, 2246–2254 (2004).
260. Miceli-Richard, C. *et al.* Significant linkage to spondyloarthropathy on 9q31-34. *Hum. Mol. Genet.* **13**, 1641–1648 (2004).
261. Lee, Y. H., Rho, Y. H., Choi, S. J., Ji, J. D. & Song, G. G. Ankylosing spondylitis susceptibility loci defined by genome-search meta-analysis. *J. Hum. Genet.* **50**, 453–459 (2005).

262. Carter, K. W. *et al.* Combined analysis of three whole genome linkage scans for Ankylosing Spondylitis. *Rheumatology (Oxford)* **46**, 763–771 (2007).
263. Hoyle, E., Laval, S. H., Calin, A., Wordsworth, B. P. & Brown, M. A. The X-chromosome and susceptibility to ankylosing spondylitis. *Arthritis Rheum.* **43**, 1353–1355 (2000).
264. Zinovieva, E. *et al.* Comprehensive linkage and association analyses identify haplotype, near to the TNFSF15 gene, significantly associated with spondyloarthritis. *PLoS Genet.* **5**, e1000528 (2009).
265. Zinovieva, E. *et al.* Lack of association between Tenascin-C gene and spondyloarthritis. *Rheumatology (Oxford)* **47**, 1655–1658 (2008).
266. Zinovieva, E. *et al.* Systematic candidate gene investigations in the SPA2 locus (9q32) show an association between TNFSF8 and susceptibility to spondylarthritis. *Arthritis Rheum.* **63**, 1853–1859 (2011).
267. Burton, P. R. *et al.* Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat. Genet.* **39**, 1329–1337 (2007).
268. Reveille, J. D. *et al.* Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat. Genet.* **42**, 123–127 (2010).
269. Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat. Genet.* **43**, 761–767 (2011).
270. Lin, Z. *et al.* A genome-wide association study in Han Chinese identifies new susceptibility loci for ankylosing spondylitis. *Nat. Genet.* **44**, 73–77 (2012).
271. Tsui, H. W., Inman, R. D., Reveille, J. D. & Tsui, F. W. L. Association of a TNAP haplotype with ankylosing spondylitis. *Arthritis Rheum.* **56**, 234–243 (2007).
272. Cheng, N. *et al.* No significant association between genetic polymorphisms in the TNAP gene and ankylosing spondylitis in the Chinese Han population. *Rheumatol. Int.* **29**, 305–310 (2009).
273. Liu, Z., Cui, Y., Zhou, X., Zhang, X. & Han, J. Association of mineralization-related genes TNAP and ANKH polymorphisms with ankylosing spondylitis in the Chinese Han population. *Biosci Trends* **7**, 89–92 (2013).
274. Zeng, Z. *et al.* Association of FCRL4 polymorphisms on disease susceptibility and severity of ankylosing spondylitis in Chinese Han population. *Clin. Rheumatol.* **31**, 1449–1454 (2012).
275. Duan, Z.-H. *et al.* The FCGR2B rs10917661 polymorphism may confer susceptibility to ankylosing spondylitis in Han Chinese: a case-control study. *Scand. J. Rheumatol.* **41**, 219–222 (2012).
276. Lv, C. *et al.* Association of Interleukin-10 gene polymorphisms with ankylosing spondylitis. *Clin Invest Med* **34**, E370 (2011).
277. Goedecke, V. *et al.* Interleukin 10 polymorphisms in ankylosing spondylitis. *Genes Immun.* **4**, 74–76 (2003).

278. Mastrovic Radoncic, K., Kaluza-Schilling, W., Cikes, N. & Märker-Hermann, E. Association of interleukin-10G microsatellite polymorphism with the susceptibility of ankylosing spondylitis. *Joint Bone Spine* **80**, 113–114 (2013).
279. McGarry, F., Neilly, J., Anderson, N., Sturrock, R. & Field, M. A polymorphism within the interleukin 1 receptor antagonist (IL-1Ra) gene is associated with ankylosing spondylitis. *Rheumatology (Oxford)* **40**, 1359–1364 (2001).
280. Van der Paardt, M. *et al.* Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms in ankylosing spondylitis. *Rheumatology (Oxford)* **41**, 1419–1423 (2002).
281. Maksymowych, W. P. *et al.* High-throughput single-nucleotide polymorphism analysis of the IL1RN locus in patients with ankylosing spondylitis by matrix-assisted laser desorption ionization-time-of-flight mass spectrometry. *Arthritis Rheum.* **48**, 2011–2018 (2003).
282. Timms, A. E. *et al.* The interleukin 1 gene cluster contains a major susceptibility locus for ankylosing spondylitis. *Am. J. Hum. Genet.* **75**, 587–595 (2004).
283. Chou, C.-T. *et al.* Replication of association of IL1 gene complex members with ankylosing spondylitis in Taiwanese Chinese. *Ann. Rheum. Dis.* **65**, 1106–1109 (2006).
284. Maksymowych, W. P. *et al.* Association of the IL1 gene cluster with susceptibility to ankylosing spondylitis: an analysis of three Canadian populations. *Arthritis Rheum.* **54**, 974–985 (2006).
285. Agrawal, S. *et al.* IL1RN\*2 allele of IL-1receptor antagonist VNTR polymorphism is associated with susceptibility to ankylosing [corrected] spondylitis in Indian patients. *Clin. Rheumatol.* **27**, 573–576 (2008).
286. Sims, A.-M. *et al.* Prospective meta-analysis of interleukin 1 gene complex polymorphisms confirms associations with ankylosing spondylitis. *Ann. Rheum. Dis.* **67**, 1305–1309 (2008).
287. Guo, Z. S. *et al.* Association of IL-1 gene complex members with ankylosing spondylitis in Chinese Han population. *Int. J. Immunogenet.* **37**, 33–37 (2010).
288. Monnet, D. *et al.* Association between the IL-1 family gene cluster and spondyloarthritis. *Ann. Rheum. Dis.* **71**, 885–890 (2012).
289. Kim, T.-H. *et al.* Interleukin 1 and nuclear factor-kappaB polymorphisms in ankylosing spondylitis in Canada and Korea. *J. Rheumatol.* **32**, 1907–1910 (2005).
290. Azizi, E. *et al.* Association of CTLA4 gene polymorphism in Iranian patients with ankylosing spondylitis. *J. Clin. Immunol.* **30**, 268–271 (2010).
291. Lee, W.-Y. *et al.* Polymorphisms of cytotoxic T lymphocyte-associated antigen-4 and cytokine genes in Taiwanese patients with ankylosing spondylitis. *Tissue Antigens* **75**, 119–126 (2010).
292. Lee, S.-H. *et al.* Association of the programmed cell death 1 (PDCD1) gene polymorphism with ankylosing spondylitis in the Korean population. *Arthritis Res. Ther.* **8**, R163 (2006).

293. Liu, X. *et al.* Programmed cell death 1 gene polymorphisms is associated with ankylosing spondylitis in Chinese Han population. *Rheumatol. Int.* **31**, 209–213 (2011).
294. Yang, Q., Liu, Y., Liu, D., Zhang, Y. & Mu, K. Association of polymorphisms in the programmed cell death 1 (PD-1) and PD-1 ligand genes with ankylosing spondylitis in a Chinese population. *Clin. Exp. Rheumatol.* **29**, 13–18 (2011).
295. Soleimanifar, N. *et al.* Study of programmed cell death 1 (PDCD1) gene polymorphisms in Iranian patients with ankylosing spondylitis. *Inflammation* **34**, 707–712 (2011).
296. Timms, A. E., Zhang, Y., Bradbury, L., Wordsworth, B. P. & Brown, M. A. Investigation of the role of ANKH in ankylosing spondylitis. *Arthritis Rheum.* **48**, 2898–2902 (2003).
297. Tsui, F. W. L. *et al.* Novel genetic markers in the 5'-flanking region of ANKH are associated with ankylosing spondylitis. *Arthritis Rheum.* **48**, 791–797 (2003).
298. Pimentel-Santos, F. M. *et al.* ANKH and susceptibility to and severity of ankylosing spondylitis. *J. Rheumatol.* **39**, 131–134 (2012).
299. Furuichi, T. *et al.* Association of the MSX2 gene polymorphisms with ankylosing spondylitis in Japanese. *J. Hum. Genet.* **53**, 419–424 (2008).
300. Chang, X. *et al.* Carbonic anhydrase I (CA1) is involved in the process of bone formation and is susceptible to ankylosing spondylitis. *Arthritis Res. Ther.* **14**, R176 (2012).
301. Liu, D., Yang, Q., Li, M., Mu, K. & Zhang, Y. Association of an asporin repeat polymorphism with ankylosing spondylitis in Han Chinese population: a case-control study. *Clin Invest Med* **33**, E63–68 (2010).
302. Van der Paardt, M. *et al.* No evidence for involvement of the Toll-like receptor 4 (TLR4) A896G and CD14-C260T polymorphisms in susceptibility to ankylosing spondylitis. *Ann. Rheum. Dis.* **64**, 235–238 (2005).
303. Gergely, P., Jr, Blazsek, A., Weiszhar, Z., Pazár, B. & Poór, G. Lack of genetic association of the Toll-like receptor 4 (TLR4) Asp299Gly and Thr399Ile polymorphisms with spondylarthropathies in a Hungarian population. *Rheumatology (Oxford)* **45**, 1194–1196 (2006).
304. Adam, R., Sturrock, R. D. & Gracie, J. A. TLR4 mutations (Asp299Gly and Thr399Ile) are not associated with ankylosing spondylitis. *Ann. Rheum. Dis.* **65**, 1099–1101 (2006).
305. Snelgrove, T. *et al.* Association of toll-like receptor 4 variants and ankylosing spondylitis: a case-control study. *J. Rheumatol.* **34**, 368–370 (2007).
306. Na, K.-S. *et al.* Analysis of single nucleotide polymorphisms in Toll-like receptor 4 shows no association with ankylosing spondylitis in a Korean population. *Rheumatol. Int.* **28**, 627–630 (2008).
307. Pointon, J. J. *et al.* Toll-like receptor 4 and CD14 polymorphisms in ankylosing spondylitis: evidence of a weak association in Finns. *J. Rheumatol.* **35**, 1609–1612 (2008).

308. Xu, W.-D., Liu, S.-S., Pan, H.-F. & Ye, D.-Q. Lack of association of TLR4 polymorphisms with susceptibility to rheumatoid arthritis and ankylosing spondylitis: a meta-analysis. *Joint Bone Spine* **79**, 566–569 (2012).
309. Liu, J. *et al.* The association of LRP5 gene polymorphisms with ankylosing spondylitis in a Chinese Han population. *J. Rheumatol.* **38**, 2616–2618 (2011).
310. Jin, L. *et al.* Lack of association of matrix metalloproteinase 3 (MMP3) genotypes with ankylosing spondylitis susceptibility and severity. *Rheumatology (Oxford)* **44**, 55–60 (2005).
311. Wei, J. C.-C. *et al.* Genetic polymorphisms of the matrix metalloproteinase-3 (MMP-3) and tissue inhibitors of matrix metalloproteinases-1 (TIMP-1) modulate the development of ankylosing spondylitis. *Ann. Rheum. Dis.* **68**, 1781–1786 (2009).
312. Pointon, J. J. *et al.* The histone demethylase JARID1A is associated with susceptibility to ankylosing spondylitis. *Genes Immun.* **12**, 395–398 (2011).
313. Chai, W. *et al.* JARID1A, JMY, and PTGER4 Polymorphisms Are Related to Ankylosing Spondylitis in Chinese Han Patients: A Case-Control Study. *PLoS ONE* **8**, e74794 (2013).
314. Wei, J. C.-C. *et al.* Association of ORAI1 haplotypes with the risk of HLA-B27 positive ankylosing spondylitis. *PLoS ONE* **6**, e20426 (2011).
315. Hung, Y.-H. *et al.* Ikb $\alpha$  promoter polymorphisms in patients with ankylosing spondylitis. *Rheumatol. Int.* **30**, 93–97 (2009).
316. Akkoc, N. *et al.* Increased prevalence of M694V in patients with ankylosing spondylitis: additional evidence for a link with familial mediterranean fever. *Arthritis Rheum.* **62**, 3059–3063 (2010).
317. Cosan, F. *et al.* Association of familial Mediterranean fever-related MEFV variations with ankylosing spondylitis. *Arthritis Rheum.* **62**, 3232–3236 (2010).
318. Yigit, S., Inanir, A., Karakus, N., Kesici, E. & Bozkurt, N. Common Mediterranean fever (MEFV) gene mutations associated with ankylosing spondylitis in Turkish population. *Dis. Markers* **33**, 113–118 (2012).
319. Miceli-Richard, C. *et al.* CARD15/NOD2 analyses in spondylarthropathy. *Arthritis Rheum.* **46**, 1405–1406 (2002).
320. Crane, A. M. *et al.* Role of NOD2 variants in spondylarthritis. *Arthritis Rheum.* **46**, 1629–1633 (2002).
321. D’Amato, M. The Crohn’s associated NOD2 3020InsC frameshift mutation does not confer susceptibility to ankylosing spondylitis. *J. Rheumatol.* **29**, 2470–2471 (2002).
322. Ferreirós-Vidal, I. *et al.* Lack of association of ankylosing spondylitis with the most common NOD2 susceptibility alleles to Crohn’s disease. *J. Rheumatol.* **30**, 102–104 (2003).
323. Van der Paardt, M. *et al.* CARD15 gene mutations are not associated with ankylosing spondylitis. *Genes Immun.* **4**, 77–78 (2003).

324. Kim, T.-H. *et al.* Analysis of CARD15 polymorphisms in Korean patients with ankylosing spondylitis reveals absence of common variants seen in western populations. *J. Rheumatol.* **31**, 1959–1961 (2004).
325. Kobak, S., Orman, M. & Berdeli, A. NOD2/CARD15 mutations in Turkish patients with ankylosing spondylitis. *Scand. J. Rheumatol.* **42**, 249–250 (2013).
326. Danoy, P. *et al.* Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with Crohn's disease. *PLoS Genet.* **6**, e1001195 (2010).
327. Chen, C., Zhang, X. & Wang, Y. Analysis of JAK2 and STAT3 polymorphisms in patients with ankylosing spondylitis in Chinese Han population. *Clin. Immunol.* **136**, 442–446 (2010).
328. Davidson, S. I. *et al.* Association of STAT3 and TNFRSF1A with ankylosing spondylitis in Han Chinese. *Ann. Rheum. Dis.* **70**, 289–292 (2011).
329. Shehab, D. K. *et al.* Association of angiotensin-converting enzyme (ACE) gene insertion-deletion polymorphism with spondylarthropathies. *J. Biomed. Sci.* **15**, 61–67 (2008).
330. Inanir, A. *et al.* Significant association between insertion/deletion polymorphism of the angiotensin-converting enzyme gene and ankylosing spondylitis. *Mol. Vis.* **18**, 2107–2113 (2012).
331. Jaakkola, E. *et al.* The effect of transforming growth factor beta1 gene polymorphisms in ankylosing spondylitis. *Rheumatology (Oxford)* **43**, 32–38 (2004).
332. Van der Paardt, M. *et al.* Susceptibility to ankylosing spondylitis: no evidence for the involvement of transforming growth factor beta 1 (TGFB1) gene polymorphisms. *Ann. Rheum. Dis.* **64**, 616–619 (2005).
333. Wu, W. *et al.* Susceptibility to ankylosing spondylitis: evidence for the role of ERAP1, TGFb1 and TLR9 gene polymorphisms. *Rheumatol. Int.* **32**, 2517–2521 (2012).
334. Beyeler, C., Armstrong, M., Bird, H. A., Idle, J. R. & Daly, A. K. Relationship between genotype for the cytochrome P450 CYP2D6 and susceptibility to ankylosing spondylitis and rheumatoid arthritis. *Ann. Rheum. Dis.* **55**, 66–68 (1996).
335. Brown, M. A. *et al.* Polymorphisms of the CYP2D6 gene increase susceptibility to ankylosing spondylitis. *Hum. Mol. Genet.* **9**, 1563–1566 (2000).
336. Davidson, S. I. *et al.* Association of ERAP1, but not IL23R, with ankylosing spondylitis in a Han Chinese population. *Arthritis Rheum.* **60**, 3263–3268 (2009).
337. Harvey, D. *et al.* Investigating the genetic association between ERAP1 and ankylosing spondylitis. *Hum. Mol. Genet.* **18**, 4204–4212 (2009).
338. Maksymowych, W. P. *et al.* Association of a specific ERAP1/ARTS1 haplotype with disease susceptibility in ankylosing spondylitis. *Arthritis Rheum.* **60**, 1317–1323 (2009).
339. Pimentel-Santos, F. M. *et al.* Association of IL23R and ERAP1 genes with ankylosing spondylitis in a Portuguese population. *Clin. Exp. Rheumatol.* **27**, 800–806 (2009).

340. Choi, C.-B. *et al.* ARTS1 polymorphisms are associated with ankylosing spondylitis in Koreans. *Ann. Rheum. Dis.* **69**, 582–584 (2010).
341. Pazár, B. *et al.* Association of ARTS1 gene polymorphisms with ankylosing spondylitis in the Hungarian population: the rs27044 variant is associated with HLA-B\*2705 subtype in Hungarian patients with ankylosing spondylitis. *J. Rheumatol.* **37**, 379–384 (2010).
342. Tsui, F. W. L. *et al.* Association of an ERAP1 ERAP2 haplotype with familial ankylosing spondylitis. *Ann. Rheum. Dis.* **69**, 733–736 (2010).
343. Bang, S.-Y. *et al.* Genetic studies of ankylosing spondylitis in Koreans confirm associations with ERAP1 and 2p15 reported in white patients. *J. Rheumatol.* **38**, 322–324 (2011).
344. Lee, Y. H., Choi, S. J., Ji, J. D. & Song, G. G. Associations between ERAP1 polymorphisms and ankylosing spondylitis susceptibility: a meta-analysis. *Inflamm. Res.* **60**, 999–1003 (2011).
345. Szczypiorska, M. *et al.* ERAP1 polymorphisms and haplotypes are associated with ankylosing spondylitis susceptibility and functional severity in a Spanish population. *Rheumatology (Oxford)* **50**, 1969–1975 (2011).
346. Chen, R., Yao, L., Meng, T. & Xu, W. The association between seven ERAP1 polymorphisms and ankylosing spondylitis susceptibility: a meta-analysis involving 8,530 cases and 12,449 controls. *Rheumatol. Int.* **32**, 909–914 (2012).
347. Saric, T. *et al.* An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* **3**, 1169–1176 (2002).
348. York, I. A. *et al.* The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat. Immunol.* **3**, 1177–1184 (2002).
349. Cui, X. *et al.* Identification of ARTS-1 as a novel TNFR1-binding protein that promotes TNFR1 ectodomain shedding. *J. Clin. Invest.* **110**, 515–526 (2002).
350. Cui, X., Rouhani, F. N., Hawari, F. & Levine, S. J. An aminopeptidase, ARTS-1, is required for interleukin-6 receptor shedding. *J. Biol. Chem.* **278**, 28677–28685 (2003).
351. Cui, X., Rouhani, F. N., Hawari, F. & Levine, S. J. Shedding of the type II IL-1 decoy receptor requires a multifunctional aminopeptidase, aminopeptidase regulator of TNF receptor type 1 shedding. *J. Immunol.* **171**, 6814–6819 (2003).
352. Wang, C.-M. *et al.* ERAP1 genetic variations associated with HLA-B27 interaction and disease severity of syndesmophytes formation in Taiwanese ankylosing spondylitis. *Arthritis Res. Ther.* **14**, R125 (2012).
353. Cherciu, M. *et al.* Functional variants of ERAP1 gene are associated with HLA-B27 positive spondyloarthritis. *Tissue Antigens* (2013). doi:10.1111/tan.12158

354. Haroon, N., Tsui, F. W. L., Chiu, B., Tsui, H. W. & Inman, R. D. Serum cytokine receptors in ankylosing spondylitis: relationship to inflammatory markers and endoplasmic reticulum aminopeptidase polymorphisms. *J. Rheumatol.* **37**, 1907–1910 (2010).
355. Parkes, M., Cortes, A., van Heel, D. A. & Brown, M. A. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat. Rev. Genet.* **14**, 661–673 (2013).
356. Goto, Y., Hattori, A., Ishii, Y. & Tsujimoto, M. Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. *FEBS Lett.* **580**, 1833–1838 (2006).
357. Evnouchidou, I. *et al.* Cutting Edge: Coding single nucleotide polymorphisms of endoplasmic reticulum aminopeptidase 1 can affect antigenic peptide generation in vitro by influencing basic enzymatic properties of the enzyme. *J. Immunol.* **186**, 1909–1913 (2011).
358. Kochan, G. *et al.* Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7745–7750 (2011).
359. Seregin, S. S. *et al.* Endoplasmic reticulum aminopeptidase-1 alleles associated with increased risk of ankylosing spondylitis reduce HLA-B27 mediated presentation of multiple antigens. *Autoimmunity* (2013). doi:10.3109/08916934.2013.819855
360. Reeves, E., Edwards, C. J., Elliott, T. & James, E. Naturally Occurring ERAP1 Haplotypes Encode Functionally Distinct Alleles with Fine Substrate Specificity. *J. Immunol.* **191**, 35–43 (2013).
361. Fruci, D. *et al.* Expression of endoplasmic reticulum aminopeptidases in EBV-B cell lines from healthy donors and in leukemia/lymphoma, carcinoma, and melanoma cell lines. *J. Immunol.* **176**, 4869–4879 (2006).
362. García-Medel, N. *et al.* Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 1 polymorphism and HLA-B27 in vivo. *Mol. Cell Proteomics* **11**, 1416–1429 (2012).
363. Martín-Esteban, A., Gómez-Molina, P., Sanz-Bravo, A. & López de Castro, J. A. Combined effects of ankylosing spondylitis-associated ERAP1 polymorphisms outside the catalytic and peptide-binding sites on the processing of natural HLA-B27 ligands. *J. Biol. Chem.* **289**, 3978–3990 (2014).
364. Rueda, B. *et al.* The IL23R Arg381Gln non-synonymous polymorphism confers susceptibility to ankylosing spondylitis. *Ann. Rheum. Dis.* **67**, 1451–1454 (2008).
365. Karaderi, T. *et al.* Association between the interleukin 23 receptor and ankylosing spondylitis is confirmed by a new UK case-control study and meta-analysis of published series. *Rheumatology (Oxford)* **48**, 386–389 (2009).
366. Sáfrány, E. *et al.* Variants of the IL23R gene are associated with ankylosing spondylitis but not with Sjögren syndrome in Hungarian population samples. *Scand. J. Immunol.* **70**, 68–74 (2009).



367. Sung, I.-H. *et al.* IL-23R polymorphisms in patients with ankylosing spondylitis in Korea. *J. Rheumatol.* **36**, 1003–1005 (2009).
368. Chen, C., Zhang, X., Li, J. & Wang, Y. Associations of IL-23R polymorphisms with ankylosing spondylitis in East Asian population: a new case-control study and a meta-analysis. *Int. J. Immunogenet.* **39**, 126–130 (2012).
369. Davidson, S. I. *et al.* Brief report: high-throughput sequencing of IL23R reveals a low-frequency, nonsynonymous single-nucleotide polymorphism that is associated with ankylosing spondylitis in a Han Chinese population. *Arthritis Rheum.* **65**, 1747–1752 (2013).
370. Dong, H., Li, Q., Zhang, Y., Tan, W. & Jiang, Z. IL23R gene confers susceptibility to ankylosing spondylitis concomitant with uveitis in a Han Chinese population. *PLoS ONE* **8**, e67505 (2013).
371. Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V. K. IL-17 and Th17 Cells. *Annu. Rev. Immunol.* **27**, 485–517 (2009).
372. Park, H. *et al.* A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat. Immunol.* **6**, 1133–1141 (2005).
373. Smith, J. A. & Colbert, R. A. Review: The interleukin-23/interleukin-17 axis in spondyloarthritis pathogenesis: Th17 and beyond. *Arthritis & Rheumatol* **66**, 231–241 (2014).
374. Wendling, D., Cedoz, J.-P., Racadot, E. & Dumoulin, G. Serum IL-17, BMP-7, and bone turnover markers in patients with ankylosing spondylitis. *Joint Bone Spine* **74**, 304–305 (2007).
375. Jandus, C. *et al.* Increased numbers of circulating polyfunctional Th17 memory cells in patients with seronegative spondylarthritides. *Arthritis Rheum.* **58**, 2307–2317 (2008).
376. Di Meglio, P. *et al.* The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PLoS ONE* **6**, e17160 (2011).
377. Pidasheva, S. *et al.* Functional studies on the IBD susceptibility gene IL23R implicate reduced receptor function in the protective genetic variant R381Q. *PLoS ONE* **6**, e25038 (2011).
378. Sarin, R., Wu, X. & Abraham, C. Inflammatory disease protective R381Q IL23 receptor polymorphism results in decreased primary CD4+ and CD8+ human T-cell functional responses. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9560–9565 (2011).
379. Coffre, M. *et al.* Combinatorial control of Th17 and Th1 cell functions by genetic variations in genes associated with the interleukin-23 signaling pathway in spondyloarthritis. *Arthritis Rheum.* **65**, 1510–1521 (2013).
380. Melzer, D. *et al.* A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet.* **4**, e1000072 (2008).
381. Vargas-Alarcón, G. *et al.* Association study of LMP gene polymorphisms in Mexican patients with spondyloarthritis. *Hum. Immunol.* **65**, 1437–1442 (2004).

382. Vargas-Alarcón, G. *et al.* Tumor necrosis factor- $\alpha$  promoter polymorphisms in Mexican patients with spondyloarthritis. *Hum. Immunol.* **67**, 826–832 (2006).
383. Kadi, A. *et al.* Investigating the genetic association between ERAP1 and spondyloarthritis. *Ann. Rheum. Dis.* **72**, 608–613 (2013).
384. Kadi, A. *et al.* Brief report: the IL23R nonsynonymous polymorphism rs11209026 is associated with radiographic sacroiliitis in spondyloarthritis. *Arthritis Rheum.* **65**, 2655–2660 (2013).
385. Poddubnyy, D. *et al.* Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. *Ann. Rheum. Dis.* **70**, 1369–1374 (2011).
386. Assassi, S. *et al.* Whole-blood gene expression profiling in ankylosing spondylitis shows upregulation of toll-like receptor 4 and 5. *J. Rheumatol.* **38**, 87–98 (2011).
387. Duan, R., Leo, P., Bradbury, L., Brown, M. A. & Thomas, G. Gene expression profiling reveals a downregulation in immune-associated genes in patients with AS. *Ann. Rheum. Dis.* **69**, 1724–1729 (2010).
388. Gu, J. *et al.* A 588-gene microarray analysis of the peripheral blood mononuclear cells of spondyloarthropathy patients. *Rheumatology (Oxford)* **41**, 759–766 (2002).
389. Gu, J. *et al.* Clues to pathogenesis of spondyloarthropathy derived from synovial fluid mononuclear cell gene expression profiles. *J. Rheumatol.* **29**, 2159–2164 (2002).
390. Gu, J. *et al.* Identification of RGS1 as a candidate biomarker for undifferentiated spondylarthritis by genome-wide expression profiling and real-time polymerase chain reaction. *Arthritis Rheum.* **60**, 3269–3279 (2009).
391. Haroon, N. *et al.* From gene expression to serum proteins: biomarker discovery in ankylosing spondylitis. *Ann. Rheum. Dis.* **69**, 297–300 (2010).
392. Laukens, D. *et al.* Altered gut transcriptome in spondyloarthropathy. *Ann. Rheum. Dis.* **65**, 1293–1300 (2006).
393. Pimentel-Santos, F. M. *et al.* Whole blood transcriptional profiling in ankylosing spondylitis identifies novel candidate genes that might contribute to the inflammatory and tissue-destructive disease aspects. *Arthritis Res. Ther.* **13**, R57 (2011).
394. Rihl, M. *et al.* Technical validation of cDNA based microarray as screening technique to identify candidate genes in synovial tissue biopsy specimens from patients with spondyloarthropathy. *Ann. Rheum. Dis.* **63**, 498–507 (2004).
395. Sharma, S. M. *et al.* Insights in to the pathogenesis of axial spondyloarthropathy based on gene expression profiles. *Arthritis Res. Ther.* **11**, R168 (2009).
396. Smith, J. A. *et al.* Gene expression analysis of macrophages derived from ankylosing spondylitis patients reveals interferon-gamma dysregulation. *Arthritis Rheum.* **58**, 1640–1649 (2008).

397. Thomas, G. P. *et al.* Expression profiling in spondyloarthropathy synovial biopsies highlights changes in expression of inflammatory genes in conjunction with tissue remodelling genes. *BMC Musculoskeletal Disord* **14**, 354 (2013).
398. Xu, L. *et al.* Changes in gene expression profiles of the hip joint ligament of patients with ankylosing spondylitis revealed by DNA chip. *Clin. Rheumatol.* **31**, 1479–1491 (2012).
399. Rihl, M. *et al.* Identification of interleukin-7 as a candidate disease mediator in spondylarthritis. *Arthritis Rheum.* **58**, 3430–3435 (2008).
400. Wright, C. *et al.* Characterization of Globin RNA Interference in Gene Expression Profiling of Whole-Blood Samples. *Clinical Chemistry* **54**, 396–405 (2008).
401. Mesko, B., Poliska, S. & Nagy, L. Gene expression profiles in peripheral blood for the diagnosis of autoimmune diseases. *Trends in Molecular Medicine* **17**, 223–233 (2011).
402. Fert, I. *et al.* Reverse interferon signature is characteristic of antigen-presenting cells in human and rat spondyloarthritis. *Arthritis & Rheumatol* **66**, 841–851 (2014).
403. Campain, A. & Yang, Y. H. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics* **11**, 408 (2010).
404. Shugart, Y. Y. *et al.* An SNP linkage scan identifies significant Crohn's disease loci on chromosomes 13q13.3 and, in Jewish families, on 1p35.2 and 3q29. *Genes Immun.* **9**, 161–167 (2008).
405. Liu, Y. *et al.* A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet.* **4**, e1000041 (2008).
406. Waetzig, G. H., Seegert, D., Rosenstiel, P., Nikolaus, S. & Schreiber, S. p38 mitogen-activated protein kinase is activated and linked to TNF-alpha signaling in inflammatory bowel disease. *J. Immunol.* **168**, 5342–5351 (2002).
407. Duerr, R. H. *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
408. Breban, M., Hammer, R. E., Richardson, J. A. & Taurog, J. D. Transfer of the inflammatory disease of HLA-B27 transgenic rats by bone marrow engraftment. *J. Exp. Med.* **178**, 1607–1616 (1993).
409. Breban, M. *et al.* T cells, but not thymic exposure to HLA-B27, are required for the inflammatory disease of HLA-B27 transgenic rats. *J. Immunol.* **156**, 794–803 (1996).
410. Hacquard-Bouder, C. *et al.* Defective costimulatory function is a striking feature of antigen-presenting cells in an HLA-B27-transgenic rat model of spondylarthropathy. *Arthritis Rheum.* **50**, 1624–1635 (2004).
411. Hacquard-Bouder, C. *et al.* Alteration of antigen-independent immunologic synapse formation between dendritic cells from HLA-B27-transgenic rats and CD4+ T cells: selective impairment of costimulatory molecule engagement by mature HLA-B27. *Arthritis Rheum.* **56**, 1478–1489 (2007).

412. Dhaenens, M. *et al.* Dendritic cells from spondylarthritis-prone HLA-B27-transgenic rats display altered cytoskeletal dynamics, class II major histocompatibility complex expression, and viability. *Arthritis Rheum.* **60**, 2622–2632 (2009).
413. Utriainen, L. *et al.* Expression of HLA-B27 causes loss of migratory dendritic cells in a rat model of spondylarthritis. *Arthritis Rheum.* **64**, 3199–3209 (2012).
414. Glatigny, S. *et al.* Proinflammatory Th17 cells are expanded and induced by dendritic cells in spondylarthritis-prone HLA-B27-transgenic rats. *Arthritis Rheum.* **64**, 110–120 (2012).
415. Talpin, A. *et al.* Monocyte-derived dendritic cells from HLA-B27+ axial spondyloarthritis (SpA) patients display altered functional capacity and deregulated gene expression. *Arthritis Res. Ther.* (2014).
416. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* **33 Suppl**, 228–237 (2003).
417. Lander, E. S. & Schork, N. J. Genetic dissection of complex traits. *Science* **265**, 2037–2048 (1994).
418. Chen, H., Malzahn, D., Balliu, B., Li, C. & Bailey, J. N. Testing genetic association with rare and common variants in family data. *Genet. Epidemiol.* **38 Suppl 1**, S37–43 (2014).
419. Saad, M. & Wijsman, E. M. Power of family-based association designs to detect rare variants in large pedigrees using imputed genotypes. *Genet. Epidemiol.* **38**, 1–9 (2014).
420. Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat. Genet.* **40**, 695–701 (2008).
421. Ott, J., Kamatani, Y. & Lathrop, M. Family-based designs for genome-wide association studies. *Nat. Rev. Genet.* **12**, 465–474 (2011).
422. Wright, V. Seronegative polyarthritis: a unified concept. *Arthritis Rheum.* **21**, 619–633 (1978).
423. Evans, D. M. & Cardon, L. R. Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. *Am. J. Hum. Genet.* **75**, 687–692 (2004).
424. Chen, M.-H. *et al.* Evaluation of Approaches to Identify Associated SNPs That Explain the Linkage Evidence in Nuclear Families with Affected Siblings. *Hum Hered* **69**, 104–119 (2010).
425. Azfer, A., Niu, J., Rogers, L. M., Adamski, F. M. & Kolattukudy, P. E. Activation of endoplasmic reticulum stress response during the development of ischemic heart disease. *Am. J. Physiol. Heart Circ. Physiol.* **291**, H1411–1420 (2006).
426. Zhang, Y., Zhang, M., Wu, J., Lei, G. & Li, H. Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking. *PLoS ONE* **7**, e48587 (2012).
427. Petit, M. M. *et al.* LHFP, a novel translocation partner gene of HMGIC in a lipoma, is a member of a new family of LHFP-like genes. *Genomics* **57**, 438–441 (1999).

428. Joshi, R. *et al.* Is there a higher genetic load of susceptibility loci in familial ankylosing spondylitis? *Arthritis Care Res (Hoboken)* **64**, 780–784 (2012).
429. Ioannidis, J. P. A., Ntzani, E. E. & Trikalinos, T. A. ‘Racial’ differences in genetic effects for complex diseases. *Nat. Genet.* **36**, 1312–1318 (2004).
430. Ward, M. M. *et al.* Clinical and immunogenetic prognostic factors for radiographic severity in ankylosing spondylitis. *Arthritis Rheum.* **61**, 859–866 (2009).
431. Kruglyak, L. & Lander, E. S. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* **57**, 439–454 (1995).
432. Middleton, F. A. *et al.* Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am. J. Hum. Genet.* **74**, 886–897 (2004).
433. John, S. *et al.* Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am. J. Hum. Genet.* **75**, 54–64 (2004).
434. Nsengimana, J., Renard, H. & Goldgar, D. Linkage analysis of complex diseases using microsatellites and single-nucleotide polymorphisms: application to alcoholism. *BMC Genet.* **6 Suppl 1**, S10 (2005).
435. Ma, Q. *et al.* Genome-wide linkage analysis for alcohol dependence: a comparison between single-nucleotide polymorphism and microsatellite marker assays. *BMC Genet.* **6 Suppl 1**, S8 (2005).
436. Gonzalez-Neira, A. *et al.* Genomewide high-density SNP linkage analysis of non-BRCA1/2 breast cancer families identifies various candidate regions and has greater power than microsatellite studies. *BMC Genomics* **8**, 299 (2007).
437. Clerget-Darpoux, F. & Elston, R. C. Are linkage analysis and the collection of family data dead? Prospects for family studies in the age of genome-wide association. *Hum. Hered.* **64**, 91–96 (2007).
438. Bailey-Wilson, J. E. & Wilson, A. F. Linkage analysis in the next-generation sequencing era. *Hum. Hered.* **72**, 228–236 (2011).
439. Jia, P. *et al.* Consensus Rules in Variant Detection from Next-Generation Sequencing Data. *PLoS ONE* **7**, e38470 (2012).
440. Bowden, D. W. *et al.* Molecular basis of a linkage peak: exome sequencing and family-based analysis identify a rare genetic variant in the ADIPOQ gene in the IRAS Family Study. *Hum. Mol. Genet.* **19**, 4112–4120 (2010).
441. Norton, N. *et al.* Exome sequencing and genome-wide linkage analysis in 17 families illustrate the complex contribution of TTN truncating variants to dilated cardiomyopathy. *Circ Cardiovasc Genet* **6**, 144–153 (2013).
442. Marchani, E. E. *et al.* Identification of rare variants from exome sequence in a large pedigree with autism. *Hum. Hered.* **74**, 153–164 (2012).

443. Jordan, C. T. *et al.* PSORS2 is due to mutations in CARD14. *Am. J. Hum. Genet.* **90**, 784–795 (2012).
444. Daw, E. W., Thompson, E. A. & Wijsman, E. M. Bias in multipoint linkage analysis arising from map misspecification. *Genet. Epidemiol.* **19**, 366–380 (2000).
445. Göring, H. H. & Terwilliger, J. D. Linkage analysis in the presence of errors I: complex-valued recombination fractions and complex phenotypes. *Am. J. Hum. Genet.* **66**, 1095–1106 (2000).
446. Göring, H. H. & Terwilliger, J. D. Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. *Am. J. Hum. Genet.* **66**, 1107–1118 (2000).
447. Göring, H. H. & Terwilliger, J. D. Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. *Am. J. Hum. Genet.* **66**, 1298–1309 (2000).
448. Lee, J. C. *et al.* A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature* **372**, 739–746 (1994).
449. McNamee, E. N., Collins, C. B., Lebsack, M. D. P. & Rivera-Nieves, J. Cell-Specific Inhibition of p38 $\alpha$  as a Therapeutic Strategy for Inflammatory Bowel Disease. *Gastroenterology* **138**, 1237–1239 (2010).
450. Coulthard, L. R. *et al.* Genetic variants within the MAP kinase signalling network and anti-TNF treatment response in rheumatoid arthritis patients. *Ann. Rheum. Dis.* **70**, 98–103 (2011).
451. Ruuska, M., Sahlberg, A. S., Colbert, R. A., Granfors, K. & Penttinen, M. A. Enhanced phosphorylation of STAT-1 is dependent on double-stranded RNA-dependent protein kinase signaling in HLA-B27-expressing U937 monocytic cells. *Arthritis Rheum.* **64**, 772–777 (2012).
452. Sahlberg, A. S. *et al.* Evidence that the p38 MAP kinase pathway is dysregulated in HLA-B27-expressing human monocytic cells: correlation with HLA-B27 misfolding. *Arthritis Rheum.* **56**, 2652–2662 (2007).
453. Costantino, F., Zeboulon-Ktorza, N., Said-Nahal, R., D’Agostino, M.-A. & Breban, M. FRI0115 Factors Associated with Radiographic Sacroiliitis in Spondyloarthritis (SPA): Results from Cross-Sectional and Longitudinal Analyses in A Cohort of Multiplex Families. *Ann Rheum Dis* **73**, 423–423 (2014).
454. Cortes, A. *et al.* Association study of genes related to bone formation and resorption and the extent of radiographic change in ankylosing spondylitis. *Ann. Rheum. Dis.* (2014). doi:10.1136/annrheumdis-2013-204835
455. Bartolomé, N. *et al.* Genetic polymorphisms inside and outside the MHC improve prediction of AS radiographic severity in addition to clinical variables. *Rheumatology (Oxford)* **51**, 1471–1478 (2012).
456. Wang, C.-M. *et al.* ERAP1 genetic variations associated with HLA-B27 interaction and disease severity of syndesmophytes formation in Taiwanese ankylosing spondylitis. *Arthritis Res. Ther.* **14**, R125 (2012).
457. Brown, M. A. *et al.* Identification of major loci controlling clinical manifestations of ankylosing spondylitis. *Arthritis Rheum.* **48**, 2234–2239 (2003).

458. Maksymowych, W. P., Adlam, N., Lind, D. & Russell, A. S. Polymorphism of the LMP2 gene and disease phenotype in ankylosing spondylitis: no association with disease severity. *Clin. Rheumatol.* **16**, 461–465 (1997).
459. Brown, M. A. *et al.* The effect of HLA-DR genes on susceptibility to and severity of ankylosing spondylitis. *Arthritis Rheum.* **41**, 460–465 (1998).
460. Jin, L. *et al.* Lack of association of matrix metalloproteinase 3 (MMP3) genotypes with ankylosing spondylitis susceptibility and severity. *Rheumatology (Oxford)* **44**, 55–60 (2005).
461. Sousa, E. *et al.* Ankylosing spondylitis susceptibility and severity--contribution of TNF gene promoter polymorphisms at positions -238 and -308. *Ann. N. Y. Acad. Sci.* **1173**, 581–588 (2009).
462. Szczypiorska, M. *et al.* ERAP1 polymorphisms and haplotypes are associated with ankylosing spondylitis susceptibility and functional severity in a Spanish population. *Rheumatology (Oxford)* **50**, 1969–1975 (2011).
463. Wong, R.-H. *et al.* Association of IL-12B genetic polymorphism with the susceptibility and disease severity of ankylosing spondylitis. *J. Rheumatol.* **39**, 135–140 (2012).
464. Schiotis, R. *et al.* Both baseline clinical factors and genetic polymorphisms influence the development of severe functional status in ankylosing spondylitis. *PLoS ONE* **7**, e43428 (2012).
465. Cooper, G. M. & Shendure, J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat. Rev. Genet.* **12**, 628–640 (2011).
466. Sadee, W. *et al.* Missing heritability of common diseases and treatments outside the protein-coding exome. *Hum. Genet.* (2014). doi:10.1007/s00439-014-1476-7
467. Kobie, J. J., Treanor, J. J. & Ritchlin, C. T. Transient decrease in human peripheral blood myeloid dendritic cells following influenza vaccination correlates with induction of serum antibody. *Immunol. Invest.* **43**, 606–615 (2014).
468. Slobodin, G. *et al.* Phenotype of resting and activated monocyte-derived dendritic cells grown from peripheral blood of patients with ankylosing spondylitis. *Inflammation* **35**, 772–775 (2012).
469. Bar-Joseph, Z., Gitter, A. & Simon, I. Studying and modelling dynamic biological processes using time-series gene expression data. *Nat. Rev. Genet.* **13**, 552–564 (2012).
470. Huang, Y. *et al.* Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. *PLoS Genet.* **7**, e1002234 (2011).
471. Menke, A. *et al.* Dexamethasone Stimulated Gene Expression in Peripheral Blood is a Sensitive Marker for Glucocorticoid Receptor Resistance in Depressed Patients. *Neuropsychopharmacology* **37**, 1455–1464 (2012).
472. Hoppe, B. *et al.* Fibrinogen and factor XIII A-subunit genotypes interactively influence C-reactive protein levels during inflammation. *Ann. Rheum. Dis.* **71**, 1163–1169 (2012).

473. Sarraj, B., Ludányi, K., Glant, T. T., Finnegan, A. & Mikecz, K. Expression of CD44 and L-selectin in the innate immune system is required for severe joint inflammation in the proteoglycan-induced murine model of rheumatoid arthritis. *J. Immunol.* **177**, 1932–1940 (2006).
474. Dancevic, C. M. *et al.* Biosynthesis and expression of a disintegrin-like and metalloproteinase domain with thrombospondin-1 repeats-15: a novel versican-cleaving proteoglycanase. *J. Biol. Chem.* **288**, 37267–37276 (2013).
475. Yokota, H., Goldring, M. B. & Sun, H. B. CITED2-mediated regulation of MMP-1 and MMP-13 in human chondrocytes under flow shear. *J. Biol. Chem.* **278**, 47275–47280 (2003).
476. Corr, M. Wnt signaling in ankylosing spondylitis. *Clin. Rheumatol.* **33**, 759–762 (2014).
477. Manicassamy, S. *et al.* Activation of beta-catenin in dendritic cells regulates immunity versus tolerance in the intestine. *Science* **329**, 849–853 (2010).
478. Cheung, V. G. *et al.* Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat. Genet.* **33**, 422–425 (2003).
479. Hampe, J. *et al.* Association between insertion mutation in NOD2 gene and Crohn’s disease in German and British populations. *Lancet* **357**, 1925–1928 (2001).
480. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nat Rev Genet* **10**, 184–194 (2009).
481. Nicolae, D. L. *et al.* Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet.* **6**, e1000888 (2010).
482. Wu, L. *et al.* Variation and genetic control of protein abundance in humans. *Nature* (2013). doi:10.1038/nature12223
483. Goto, Y., Ogawa, K., Hattori, A. & Tsujimoto, M. Secretion of endoplasmic reticulum aminopeptidase 1 is involved in the activation of macrophages induced by lipopolysaccharide and interferon-gamma. *J. Biol. Chem.* **286**, 21906–21914 (2011).



## **ANNEXES**

## **Annexe 1**

---

Article publié dans « Annals of the Rheumatic Diseases » en décembre 2013 (online first).

### **Prevalence of Spondyloarthritis in Reference to HLA-B27 in the French Population:**

#### **Results of the GAZEL Cohort**

Félicie Costantino<sup>1,2\*</sup>, Alice Talpin<sup>1,2\*</sup>, Roula Said-Nahal<sup>3</sup>, Marcel Goldberg<sup>4,5</sup>, Joseph Henny<sup>4,5</sup>, Gilles Chiocchia<sup>1,2,5</sup>, Henri-Jean Garchon<sup>1,2,5,6</sup>, Marie Zins<sup>4,5</sup>, Maxime Breban<sup>1,2,3,5</sup>

*1 – Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Paris, 75014, France;*

*2 – INSERM U1016 - Paris, 75014, France;*

*3 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, 92100, France;*

*4 – Population-Based Epidemiological Cohorts, Research Platform, UMRS 1018, Centre for Research in Epidemiology and Population Health, Villejuif, France;*

*5 – UFR des Sciences de la Santé, Simone Veil, Versailles Saint Quentin en Yvelines University, Montigny-Le-Bretonneux, 78180, France;*

*6 – Genetics Division, Ambroise Paré Hospital (AP-HP) Boulogne-Billancourt, 92100, France;*

*\* The two first authors contributed equally to that work.*

Address reprint requests and correspondence to:

Maxime Breban, M.D., Ph.D., Rheumatology Division

Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France

Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865

E-mail: maxime.breban@apr.aphp.fr

Keywords: ankylosing spondylitis; spondyloarthritis; prevalence; HLA-B27; rs4349859.

Word Count: 2,604

## **ABSTRACT**

**Objective:** To estimate the prevalence of SpA in reference to HLA-B27 in the French population.

**Methods:** In 1989, 20,625 employees of the French national gas and electricity company aged 35-50 years were enrolled in the GAZEL cohort. In 2010, 18,757 still active participants were screened by a questionnaire validated for the detection of SpA. Responders with available DNA were retained for further studies. Pelvic radiograph and HLA-B27 typing were performed in all the self-reported cases of SpA or psoriatic arthritis. Self-reported diagnosis was verified by qualified rheumatologist. HLA-B27 determination was also performed in subjects without any SpA feature.

**Results:** The target population consisted of 6,556 responders with available DNA. Their male:female ratio was 3.6 and their mean age was  $65.5 \pm 3.3$  years. A diagnosis of SpA was confirmed in 32 of the 72 self-reported cases, 75% of them being HLA-B27 positive. Estimated SpA prevalence adjusted for sex was 0.43% (95%CI: 0.26 to 0.70%). HLA-B27 positivity rate in 2,466 healthy controls was 6.9% (95%CI: 5.9 to 7.9%). The relative risk of SpA in HLA-B27+ individuals was 39 (95%CI: 17 to 86).

**Conclusion:** We estimated the prevalence of SpA in the French population in 2010 to 0.43%. With an estimated prevalence of 75.0% in SpA and 6.9% in healthy controls, HLA-B27 increased the disease-risk by 39 fold, as compared to HLA-B27 negative subjects.

## INTRODUCTION

Spondyloarthritis (SpA) is a frequent chronic inflammatory disorder. Depending on its clinico-radiological presentation, it has classically been subdivided into the following subsets: ankylosing spondylitis (AS), psoriatic arthritis (PsA), arthritis associated with IBD (AIBD), reactive arthritis (ReA) and undifferentiated SpA (uSpA) [1]. More recently, the Assessment of SpondyloArthritis international Society (ASAS) developed new classification criteria based on the anatomic distribution of rheumatic manifestations, subdividing the disease entity into axial and peripheral SpA [2,3].

Until now, most epidemiological studies in the field of SpA have been focused on AS. Recently, several studies have embraced all the SpA subtypes together [4] but only one of them has taken into account the ASAS classification criteria [5]. One of the goals of these criteria was to allow the conduct of clinical trials for the treatment of non-radiographic SpA. Several studies using them have already been published and showed response rates to TNF blockers similar to those initially reported in AS [6,7]. Thus, it seems important to obtain prevalence data according to ASAS classification criteria with regard to public health.

HLA-B27 is a major genetic factor of susceptibility for SpA. Its population prevalence has been previously estimated in many countries, leading to infer the proportion of HLA-B27+ individuals expected to develop SpA [8,9], but to our best knowledge, such proportion has not been directly established in previous studies. HLA-B27 is one of few biologic tests that may help with SpA diagnosis and is a key component of internationally validated classification criteria for SpA [2,3]. Thus, it is important to determine its prevalence as precisely as possible to appreciate the diagnostic performance of the test [10].

The aim of the present work was to estimate the population prevalence of SpA in reference to HLA-B27 in a cohort of French workers, the GAZEL cohort.

## SUBJECTS AND METHODS

### *Subjects*

The GAZEL cohort is an ongoing general population-based cohort established among the employees of the French national electricity and gas company (EDF-GDF) [11]. This firm employs approximately 150,000 people of diversified socio-economic statuses throughout France. The epidemiologic profile of this population is very close to that of the French general population [12]. Noteworthy, the staff of this company is very stable because they benefit of

protective social advantages. In particular, employees of the company usually start their career in the twenties and remain hired by the company, even if they become invalid or unable to work because of chronic illness. These conditions minimized the risk of left-censoring effect in the cohort (i.e. the risk of missing cases because their chronic illness would keep them away from workforce). In January 1989, 44,922 persons (31,411 male employees aged 40-50 years and 13,511 female employees aged 35-50 years) were invited to enroll in the cohort, of whom 20,625 accepted (15,011 men and 5,614 women). Each year, data on health, lifestyle, socio-economic and occupational factors were collected via a mail questionnaire. Blood samples or buccal cells were stored for approximately one third of the participants. In 2010, the 18,757 remaining participants were invited to answer to a medical questionnaire previously validated for the detection of SpA [13]. All the responders with stored biological sample available for DNA extraction were kept for further analysis. This subset of subjects with stored biological sample was representative of the whole cohort with regard to most determinants [14]. The study flow chart is shown in **Figure 1**.

The GAZEL protocol was approved by the French authority for data confidentiality (“Commission Nationale Informatique et Liberté”) and by the Ethics Evaluation Committee of the “Institut National de la Santé et de la Recherche Médicale (INSERM)” (IRB0000388, FWA00005831). Informed consent for this specific study was obtained from each participant and was approved by the local ethics committees of Ambroise Paré Hospital (Boulogne-Billancourt, France).

### ***Screening and validation of SpA diagnosis***

The screening questionnaire researched previous diagnosis of SpA or PsA as well as suggestive features of SpA as previously described [13]. Performance of self-reported diagnosis of SpA, as previously reported in the French population for an estimated prevalence of the disease of 0.5% was close to 100% for negative predictive value and low for positive predictive value (9%) [15].

This is why patients with self-reported diagnosis of SpA or PsA were interviewed by phone by a qualified rheumatologist (RSN) to verify clinical grounds of the diagnosis. For all the self-reported cases of SpA or PsA, antero-posterior radiograph of the pelvis and HLA-B27 typing were performed. Using available information, a diagnosis of SpA was confirmed or dismissed by RSN, in reference to ASAS classification criteria for axial and peripheral SpA. Peripheral PsA cases who did not fulfill ASAS classification criteria for peripheral SpA (i.e. patients suffering of peripheral joint arthritis of symmetric distribution and predominating in

the upper limbs, in a context of psoriasis but without other feature of SpA) were classified separately.

All the subjects with available biological samples who had answered negatively to all the questionnaire's items were considered as definite healthy controls from the perspective of SpA. Subjects who did not report SpA or PsA but who had answered positively to at least one questionnaire's item were not investigated further.

### ***Determination of HLA-B27 status***

HLA-B27 status was determined in the suspected SpA cases and in the definite healthy controls. Two methods were used for this determination: allele-specific PCR and genotyping of the single-nucleotide polymorphism (SNP) rs4349859, known to strongly tag HLA-B27 allele [16].

### ***DNA extraction***

DNA was extracted from stored buccal cells (n=927) or blood samples (n=1,694). Buccal cells had been collected with a sterile cytobrush (Histobrush, Hardwood Products Company, USA) by twirling it on the inner cheek. Blood sample had been obtained by venipuncture and buffy coat was then isolated and stored. DNA was extracted with QIAamp DNA Blood Midi and Mini kit (QIAGEN, Valencia, CA) from buffy coats and cytobrushes respectively.

### ***HLA-B27 status determination***

Exons 2 and 3 of the HLA-B locus were amplified with HLA-B27 specific primers (Forward: GGTCTCACACCCTCCAGAAT, Reverse: GCGTCTCCTTCCCGTTCTCA) together with GAPDH primers (Forward: ACCACAGTCCATGCCATCA, Reverse: CCACCACCCTGTTGCTGT) as internal control. Detection of the amplified products was carried out by mean of 2% agarose gel electrophoresis.

Genotyping of rs4349859 was performed using Taqman genotyping assay (Applied Biosystems, Foster City, California, USA) according to manufacturer's recommendations.

### ***Statistical methods***

Continuous variables were expressed as mean  $\pm$  standard deviation. Crude and sex-adjusted prevalence were estimated. Adjusted prevalence was calculated based on the sex distribution in the source population as determined by the last national census (2010) [17]. The 95% confidence intervals (95%CI) were constructed using the logit transformation [18].

To evaluate SpA prevalence according to HLA-B27 status, we estimated the number of HLA-B27 positive subjects based on HLA-B27 typing in healthy controls and individuals having reported SpA. Categorical variables were compared using the chi-square test and continuous variables using the Student's T test. The level of type I error used to determine statistical significance was 5%. Statistical analyses were performed with R for Windows, version 2.15.1.

## RESULTS

### *Characteristics of the population*

As shown in **Figure 1**, 18,757 subjects (91% of the initial cohort) were screened for participation by a dedicated questionnaire. The overall response rate was 74.5%. Among the responders, 6,556 had stored buccal cells or buffy coat sample available for DNA extraction, representing 34.9% of the screened cohort. They constituted the target population for the rest of the study. Participants included in this study were not different from the whole GAZEL cohort, except for being more often retired and having a higher socio-economical level.

### *SpA prevalence*

Seventy-two subjects declared to have AS, SpA or PsA. As shown in **Figure 2**, a diagnosis of SpA was confirmed in 32 of them, a diagnosis of peripheral PsA in 7 of them, and another diagnosis in 28 of them. In 2 cases, diagnosis remained doubtful. Clinical characteristics of the confirmed cases are summarized in **Table 1** and presented in details in Supplementary table 1. All the confirmed cases fulfilled ASAS classification criteria for SpA and 75.0% of them were HLA-B27 positive (80% of the AS cases and 67% of the other SpA cases).

The overall sex-adjusted estimated prevalence of SpA was 0.43% (95%CI: 0.26 to 0.70%). **Table 2** shows the prevalence of SpA, according to age, sex and HLA-B27 status. There was no significant difference in prevalence according to age or sex.

### *HLA-B27 prevalence in healthy controls*

Individuals having answered negatively to all items of the arthritis questionnaire (n = 2,551, i.e. 38.9% of the target population) were considered as definite healthy controls and used for HLA-B27 status determination, by allele-specific PCR.

Overall, 6.9% (95%CI: 5.9 to 7.9%) of the tested samples were HLA-B27 positive. There was no significant difference between males and females. However, a trend towards a

lower rate of HLA-B27 was observed in subjects older than 70 years (5.1%; 95%CI: 3.4 to 7.4%), as compared to younger subjects (7.3%; 95%CI: 6.2 to 8.6%), that was close to significance ( $p=0.06$ ) (**Table 3**). The relative risk of developing SpA in HLA-B27 positive individuals, as compared to HLA-B27 negative ones was 39 (95%CI: 17 to 86).

The minor allele frequency of rs4349859 (corresponding to allele A which tags HLA-B27) was 6.8% (95%CI: 5.9 to 7.9%). In comparison to HLA-B27 PCR results, rs4349859 showed high specificity (99.7%) and sensitivity (97.0%) for HLA-B27 detection.

## DISCUSSION

Accurate public health planning requires reliable epidemiologic data to be effective. Although several recent studies have estimated the prevalence of SpA [4], as far as we know, the present study is the first to estimate both SpA and HLA-B27 prevalence in a single large-scale cohort. Moreover, to our knowledge, it is the first to report the prevalence of SpA in the general population, according to ASAS classification criteria, since *Strand et al.* studied SpA prevalence in US rheumatology practices [5]. By screening subjects from the French GAZEL cohort of more than 56 years old, we estimated SpA prevalence around 0.43% and that of HLA-B27 carriers around 6.9% in healthy controls and 75.0% in SpA patients.

Such SpA prevalence appears higher than the 0.3% value reported by *Saraux et al.* from a survey conducted in the French adult population, in 2001 [19], albeit the 95%CI of estimate overlaps between both studies. Such discrepancy can be explained by the fact that the age of subjects investigated here was  $\geq 57$  years. Given that a diagnosis of SpA is most of the time made before the age of 50 years [20], one can assume that the present study yielded a close approximate of the cumulative risk of developing SpA in the French population. Consistent with other recently published studies, we observed no significant variation in prevalence according to sex [19,21]. AS was by far the most common subtype of SpA in our study (69%), which may appear different from other studies, in which PsA and/or uSpA were found more frequent than AS [9,19,22]. Two major factors could explain such apparent discordance. First, pelvic radiograph was obtained in SpA cases in a systematic fashion only in our study. Secondly, subjects in the present study were older than in previous ones at the time of ascertainment and some of those presenting at first with uSpA or PsA could have progressed to AS over time [23,24].

To estimate SpA prevalence, it is important to target a general population sample and not only patients referred to rheumatologist. This is why we used the GAZEL cohort, which is composed of subjects from of all French regions and of various socio-economic statuses.



Several studies have shown that the epidemiologic profile of the cohort is close to that of the French general population in terms of distribution of disorders and of causes of death [25,26]. Nevertheless some bias could be seen as a limitation to generalization of those results to the entire French population. First, there was a high proportion of men in the cohort. Second, all the subjects were more than 55 years old. Third, GAZEL participants were volunteers, thus subject to a selection effect potentially associated with their health status [27]. Fourth, French nationality is required to be employed in EDF-GDF, thus ethnic variability in the cohort was lower than in the general population.

However, such potential biases should be very limited: the prevalence rate of SpA was similar between men and women and we calculated sex-adjusted prevalence. Although elderly onset of SpA is possible [28], we can consider that very few cases should occur later given the age of the population studied and that we have estimated here the cumulative prevalence of SpA. Thus we did not adjust prevalence for age. It is unlikely that having a SpA could introduce a selection bias in EDF-GDF recruitment. Indeed, mean age at recruitment in the company was around 20 years and once recruited, all the individuals remained in the company and were proposed to enter the cohort, even in case of work disability or long-term illness. Moreover, in previous studies having assessed confounding factors in the GAZEL cohort, rheumatic condition was not associated with initial participation nor with attrition during follow-up [12,29]. All these data support that the GAZEL cohort can be considered as representative of the French Caucasian population.

As another possible source of bias, our screening was based on self-reported diagnosis so that we might have missed some undiagnosed cases, in particular in those individuals who had declared some SpA features. Again, such bias should be minor, as previously reported negative predictive value of self-reported SpA being close to 100% [15]. Self-reported diagnoses were validated by phone interview by a qualified rheumatologist. Previous reports suggested that telephone survey may represent a reliable alternative to face to face interview [15,19] with no major differences in the responses obtained with the two methods [30].

In the present study, the estimated HLA-B27 prevalence in the French population was 6.9% in controls, a result close to that of a recent study reporting a HLA-B27 prevalence of 7.5% in non-Hispanic white subjects, in US [31]. We inferred a 39-fold relative risk of developing SpA in HLA-B27 positive, as compared to B27 negative individuals. *Braun et al.* reported a lower relative risk in HLA-B27 individuals from the Berlin population (20.7) but the SpA prevalence used to estimate it (1.9%) was probably overrated in their initial study [32].

Interestingly, a trend towards lower rate of HLA-B27 positivity was found in subjects older than 70 years. In our study, the lowest rate of B27 was observed in persons born during the Second World War. Since many epidemics have occurred during this period in France, it may suggest a higher susceptibility of HLA-B27 positive individuals to tuberculosis or other unidentified infectious disease [33]. This observation is reminiscent of that recently reported in the US population by *Reveille et al* [31] in subjects aged 50-69 years. The shift observed in age range may reflect differences in pathogens exposure between France and USA.

We assessed the presence of HLA-B27 by direct PCR genotyping and also by genotyping rs4349859, a surrogate marker for B27 [16]. Our study showed good agreement between both methods. Thus, we confirmed that genotyping rs4349859 is a highly specific method to detect HLA-B27 that can be used to screen large population samples, in particular for epidemiological studies.

In conclusion, cumulative SpA prevalence in the French population in 2010 was slightly higher than that previously reported in France in 2001, with a high proportion of AS. Although less than 5% of HLA-B27 carriers have developed SpA, disease relative risk conferred by this antigen was high with a 39-fold increase. The trend towards decreased HLA-B27 rate in older subjects could also suggest other role for this molecule with a possible higher childhood-related mortality in HLA-B27 carriers due to exposure to infectious peril having vanished since the 50s in France.

## **ACKNOWLEDGMENTS AND AFFILIATIONS**

We thank Dr. Joëlle Benessiano and the Hospital Bichat-Claude Bernard Biological Resource Center (CRB-BCB, Paris, France) for DNA extraction and storage of GAZEL samples.

This work was supported by grant from Agence Nationale de la Recherche (grant ANR 2010 GEMISA). Félicie Costantino was supported by a grant from Société Française de Rhumatologie (SFR). Alice Talpin was supported by a grant from Arthritis Fondation Courtin. The GAZEL Cohort Study was funded by EDF-GDF and INSERM, and received grants from the ‘Cohortes Santé TGIR Program’, Agence nationale de la recherche (ANR) and Agence française de sécurité sanitaire de l’environnement et du travail (AFSSET).

## REFERENCES

1. Baeten D, Breban M, Lories R, et al. Are spondylarthritides related but distinct conditions or a single disease with a heterogeneous phenotype? *Arthritis Rheum* 2013;65:12–20.
2. Rudwaleit M, van der Heijde D, Landewé R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.
3. Rudwaleit M, van der Heijde D, Landewé R, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
4. Stolwijk C, Boonen A, van Tubergen A, et al. Epidemiology of spondyloarthritis. *Rheum Dis Clin North Am* 2012;38:441–76.
5. Strand V, Rao SA, Shillington AC, et al. Prevalence of axial SpA in US rheumatology practices: Assessment of ASAS criteria vs. rheumatology expert clinical diagnosis. *Arthritis Care Res (Hoboken)* Published Online First: 22 February 2013.
6. Sieper J, van der Heijde D, Dougados M, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). *Ann Rheum Dis* 2013;72:815–22.
7. Song I-H, Weiß A, Hermann K-GA, et al. Similar response rates in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis after 1 year of treatment with etanercept: results from the ESTHER trial. *Ann Rheum Dis* 2013;72:823–5.
8. Van der Linden S, Valkenburg H, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals: a family and population study. *Br J Rheumatol* 1983;22:18–9.
9. Braun J, Bollow M, Remlinger G, et al. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. *Arthritis Rheum* 1998;41:58–67.
10. Van den Bosch F, Elewaut D. Spondyloarthritis: A conundrum: classifying patients with SpA in daily practice. *Nat Rev Rheumatol* 2012;8:699–700.
11. Goldberg M, Leclerc A, Bonenfant S, et al. Cohort profile: the GAZEL Cohort Study. *Int J Epidemiol* 2007;36:32–9.
12. Goldberg M, Chastang JF, Leclerc A, et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. *Am J Epidemiol* 2001;154:373–84.

13. Said-Nahal R, Miceli-Richard C, D'Agostino MA, et al. Phenotypic diversity is not determined by independent genetic factors in familial spondylarthropathy. *Arthritis Rheum* 2001;45:478–84.
14. Zins M, Ozguler A, Bonenfant S, et al. [Constitution of a bank of biological material in the French Gazel cohort: logistical and practical aspects]. *Rev Epidemiol Sante Publique* 2003;51:143–6.
15. Guillemin F, Saraux A, Fardellone P, et al. Detection of cases of inflammatory rheumatic disorders: performance of a telephone questionnaire designed for use by patient interviewers. *Ann Rheum Dis* 2003;62:957–63.
16. Evans DM, Spencer CCA, Pointon JJ, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–7.
17. INSEE. <http://www.insee.fr/fr/> (accessed 8 Aug2013).
18. Newcombe R. Logit Confidence Intervals and the Inverse Sinh Transformation. *The American Statistician* 2001;55:200–2.
19. Saraux A, Guillemin F, Guggenbuhl P, et al. Prevalence of spondyloarthropathies in France: 2001. *Ann Rheum Dis* 2005;64:1431–5.
20. Feldtkeller E, Khan MA, van der Heijde D, et al. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatol Int* 2003;23:61–6.
21. Reveille JD, Witter JP, Weisman MH. Prevalence of axial spondylarthritis in the United States: estimates from a cross-sectional survey. *Arthritis Care Res (Hoboken)* 2012;64:905–10.
22. Haglund E, Bremander AB, Petersson IF, et al. Prevalence of spondyloarthritis and its subtypes in southern Sweden. *Ann Rheum Dis* 2011;70:943–8.
23. Mau W, Zeidler H, Mau R, et al. Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup. *J Rheumatol* 1988;15:1109–14.
24. Hanly JG, Russell ML, Gladman DD. Psoriatic spondyloarthropathy: a long term prospective study. *Ann Rheum Dis* 1988;47:386–93.
25. Chevalier A, Leclerc A, Blanc C, et al. Disparités sociales et professionnelles de la mortalité des travailleurs d'Electricité de France-Gaz de France. *Population* 1987;6:863–79.
26. Chevalier A, Goldberg M, Godard C, et al. [Cancer incidence among active male workers at Electricite de France-Gaz de France]. *Rev Epidemiol Sante Publique* 1996;44:25–36.

27. Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol* 1977;106:184–7.
28. Olivieri I, D'Angelo S, Palazzi C, et al. Late-onset spondyloarthritis: subset that should not be forgotten. *J Rheumatol* 2012;39:1110–2.
29. Goldberg M, Chastang JF, Zins M, et al. Health problems were the strongest predictors of attrition during follow-up of the GAZEL cohort. *J Clin Epidemiol* 2006;59:1213–21.
30. Siemiatycki J. A comparison of mail, telephone, and home interview strategies for household health surveys. *Am J Public Health* 1979;69:238–45.
31. Reveille JD, Hirsch R, Dillon CF, et al. The prevalence of HLA-B27 in the US: data from the US National Health and Nutrition Examination Survey, 2009. *Arthritis Rheum* 2012;64:1407–11.
32. Akkoc N, Khan MA. Overestimation of the prevalence of ankylosing spondylitis in the Berlin study: comment on the article by Braun et al. *Arthritis Rheum* 2005;52:4048–4049; author reply 4049–4050.
33. Zervas J, Constantopoulos C, Toubis M, et al. HLA-A and B antigens and pulmonary tuberculosis in Greeks. *Br J Dis Chest* 1987;81:147–9.

## TABLES AND FIGURES

**Table 1.** Clinical characteristics of confirmed SpA cases

Characteristic	SpA (n=32)
Age in years, mean $\pm$ SD	67.0 $\pm$ 3.5
Disease duration in years, mean $\pm$ SD	31.0 $\pm$ 11.1
Sex distribution: cases/total under study	
- Men	27/5,140
- Women	5/1,416
HLA-B27 positivity, %	75.0
Axial manifestations	
- Inflammatory back pain, %	75.0
- Radiographic sacroiliitis *, %	66.7
Peripheral manifestations	
- Peripheral arthritis, %	50.0
- Peripheral enthesitis, %	59.4
Extra-articular manifestations	
- Uveitis, %	15.6
- Psoriasis, %	40.6
- Inflammatory bowel disease, %	12.5
ASAS classification criteria fulfillment	
- ASAS axial SpA, %	75.0
- ASAS peripheral SpA, %	25.0
SpA subtype**	
- AS, %	66.7
- uSpA, %	13.3
- PsA, %	16.7
- AIBD, %	3.3

The registered manifestations correspond to those present at the time of investigation, or retrieved from past-medical history.

Inflammatory bowel disease: Crohn's disease or ulcerative colitis; ASAS, Assessment of SpondyloArthritis international Society; AS, ankylosing spondylitis; uSpA, undifferentiated spondyloarthritis; PsA, psoriatic arthritis; AIBD, Inflammatory bowel disease-associated arthritis;

\* Refers to radiographic sacroiliitis  $\geq$  grade II bilateral or grade III unilateral in 30 confirmed SpA cases with available pelvic X-ray.

\*\* Subtype was determined in 30 confirmed SpA cases with available pelvic X-ray

**Table 2.** Prevalence of SpA in the GAZEL cohort in 2010.

Target population	Sample, n	SpA cases, n	Prevalence, % (95%CI)
All population			
- Crude	6,556	32	0.48 (0.34 - 0.69)
- Sex-adjusted			0.43 (0.26 - 0.70)
Age group			
- 57 - 64 years	2,797	9	0.32 (0.17 - 0.62)
- ≥ 65 years	3,759	23	0.61 (0.41 - 0.92)
Sex			
- Male	5,140	27	0.53 (0.36 - 0.77)
- Female	1,416	5	0.35 (0.11 - 0.85)
HLA-B27			
- Negative	6,084*	8	0.13 (0.07 - 0.26)
- Positive	472*	24	5.08 (3.43 - 7.47)
SpA subtype			
- AS	6,556	20	0.31 (0.20 - 0.47)
- PsA	6,556	5	0.08 (0.03 - 0.18)
- uSpA	6,556	4	0.06 (0.02 - 0.16)
- AIBD	6,556	1	0.02 (0.002-0.11)

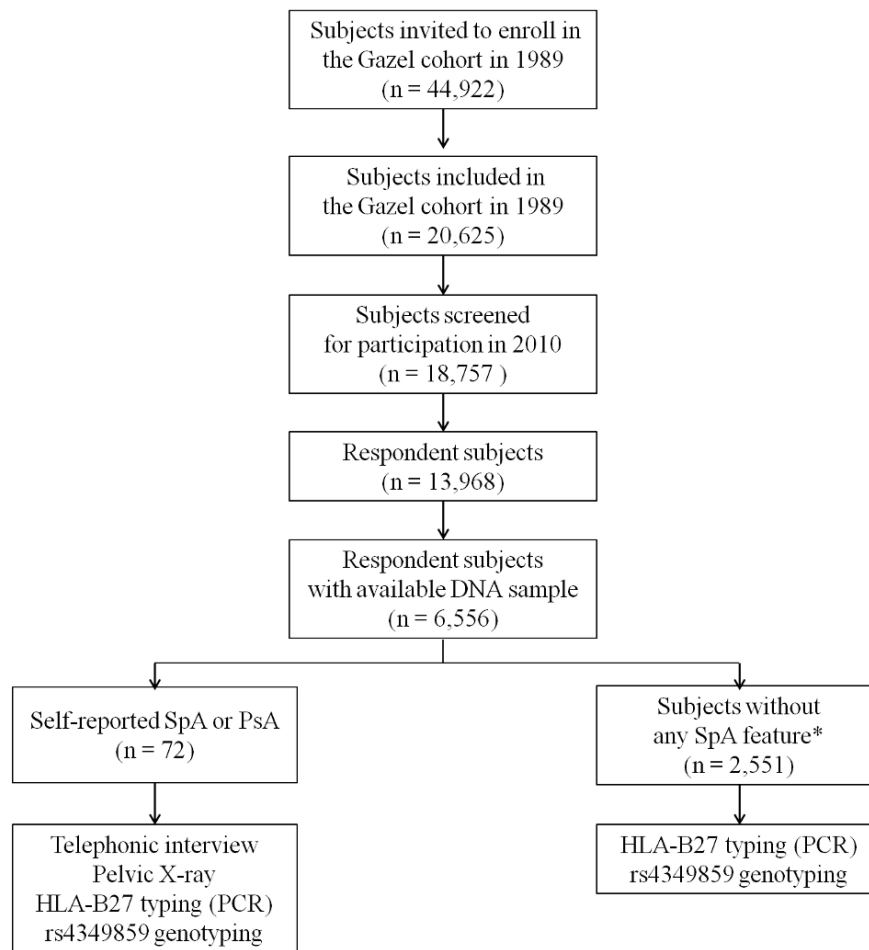
\* Estimation based on HLA-B27 typing in a sample of healthy controls (n = 2,466) and in individuals having reported SpA (n = 72)

**Table 3.** Prevalence of HLA-B27 in healthy controls of the GAZEL cohort in 2010.

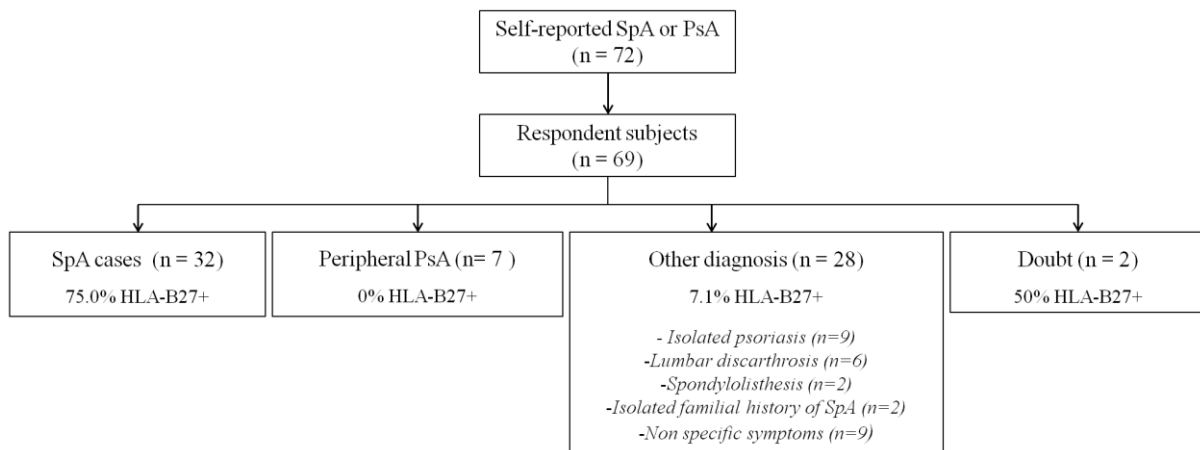
Group	Sample, n	HLA-B27+ subjects, n	Prevalence, % (95% CI)
All	2,466	169	6.9 (5.9 - 7.9)
Age group			
- 57 – 64 years	731	55	7.5 (5.8 - 9.7)
- 65 – 69 years	1,202	87	7.2 (5.9 - 8.9)
- ≥ 70 years	533	27	5.1 (3.5 - 7.3)
Sex			
- Male	2,040	134	6.6 (5.6 - 7.7)
- Female	426	35	8.2 (5.9 - 11.2)



**Figure 1.** Study flow chart



**Figure 2.** Diagnosis flow chart



**Supplementary table 1.** Detailed clinical characteristics of confirmed SpA and peripheral PsA cases

Patient	Sex	Age	Age at onset	IBP	Enthesitis	Arthritis	Uveitis	Psoriasis	IBD	Radiographic sacro-iliitis	HLA-B27	SpA subtype
<i>Axial SpA</i>												
1	Male	73	16	Yes	Yes	No	No	No	No	Grade IV	Positive	AS
2	Male	73	18	Yes	No	Yes	Yes	No	No	Grade IV	Positive	AS
3	Male	64	22	Yes	Yes	No	Yes	No	No	Grade IV	Positive	AS
4	Male	64	29	Yes	Yes	Yes	Yes	No	No	Grade IV	Positive	AS
5	Female	70	32	Yes	No	No	Yes	No	No	Grade IV	Positive	AS
6	Male	64	25	Yes	Yes	No	No	No	No	Grade IV	Positive	AS
7	Male	70	32	Yes	No	No	No	No	No	Grade IV	Positive	AS
8	Male	67	30	Yes	Yes	No	No	No	No	Grade IV	Positive	AS
9	Male	69	25	Yes	Yes	No	No	No	No	Grade IIb	Positive	AS
10	Male	66	28	Yes	No	No	No	No	No	Grade IIb	Positive	AS
11	Male	70	25	Yes	No	No	No	No	No	Grade IV	Negative	AS
12	Male	72	30	Yes	Yes	No	No	Yes	No	Grade III	Positive	AS
13	Male	68	15	Yes	No	Yes	No	Yes	No	Grade IV	Positive	AS
14	Male	65	25	Yes	Yes	No	No	Yes	No	Grade IV	Positive	AS
15	Male	70	43	Yes	No	Yes	No	Yes	Yes	Grade III	Positive	AS
16	Male	68	25	Yes	Yes	Yes	No	Yes	Yes	Grade IV	Negative	AS
17	Male	65	52	Yes	No	No	No	No	Yes	Grade IIb	Negative	AS
18	Male	65	20	Yes	No	No	No	No	No	No sacro-iliitis	Positive	uSpA
19	Male	68	26	Yes	Yes	No	No	No	No	No sacro-iliitis	Positive	uSpA
20	Male	73	30	Yes	Yes	No	No	No	No	No sacro-iliitis	Positive	uSpA
21	Female	59	30	Yes	Yes	No	No	No	No	No sacro-iliitis	Positive	uSpA
22	Male	66	50	Yes	Yes	Yes	No	Yes	No	NR	Positive	ND
23	Male	64	40	Yes	No	Yes	No	Yes	No	No sacro-iliitis	Negative	PsA
24	Female	60	35	Yes	Yes	Yes	Yes	No	Yes	No sacro-iliitis	Negative	AIBD

Patient	Sex	Age	Age at onset	IBP	Enthesitis	Arthritis	Uveitis	Psoriasis	IBD	Radiographic sacro-iliitis	HLA-B27	SpA subtype
<i>Peripheral SpA</i>												
25	Male	67	40	No	No	Yes	No	Yes	No	Grade III	Positive	AS
26	Male	64	50	No	No	Yes	No	Yes	No	Grade IIb	Negative	AS
27	Female	70	61	No	Yes	Yes	No	No	No	Grade III	Positive	AS
28	Male	65	36	No	Yes	Yes	No	Yes	No	No sacro-iliitis	Positive	PsA
29	Male	64	30	No	No	Yes	No	Yes	No	No sacro-iliitis	Negative	PsA
30	Male	64	36	No	Yes	Yes	No	Yes	No	No sacro-iliitis	Negative	PsA
31	Female	70	20	No	Yes	Yes	No	No	No	NR	Positive	ND
32	Male	67	17	No	Yes	Yes	No	Yes	No	No sacro-iliitis	Positive	PsA
<i>Peripheral PsA</i>												
1	Male	64	50	No	No	Symmetric	No	Yes	No	No sacro-iliitis	Negative	Peripheral PsA
2	Male	69	55	No	No	Symmetric	No	Yes	No	NR	Negative	Peripheral PsA
3	Male	64	60	No	No	Symmetric	No	Yes	No	NR	Negative	Peripheral PsA
4	Male	68	32	No	No	Symmetric	No	Yes	No	No sacro-iliitis	Negative	Peripheral PsA
5	Male	72	45	No	No	Symmetric	No	Yes	No	NR	Negative	Peripheral PsA
6	Male	68	47	No	No	Symmetric	No	Yes	No	No sacro-iliitis	Negative	Peripheral PsA
7	Male	65	55	No	Yes	Symmetric	No	Yes	No	No sacro-iliitis	Negative	Peripheral PsA

The registered manifestations correspond to those present at the time of investigation, or retrieved from past-medical history.

IBP: inflammatory back pain; IBD: Inflammatory bowel disease; AS, ankylosing spondylitis; uSpA, undifferentiated spondyloarthritis; PsA, psoriatic arthritis; AIBD, Inflammatory bowel disease-associated arthritis; NR: not realized; ND: not determined

## **Annexe 2**

---

Abstract accepté en communication affichée au congrès européen de rhumatologie (EULAR) 2014 et en communications orales au 9<sup>e</sup> congrès international de la spondyloarthrite en 2014 et au 27<sup>e</sup> congrès de la Société Française de Rhumatologie 2014

### **Factors associated with radiographic sacroiliitis in spondyloarthritis (SpA): results from cross-sectional and longitudinal analyses in a cohort of multiplex families.**

Félicie Costantino<sup>1,2,3</sup>, Nadine Zeboulon-Ktorza<sup>1</sup>, Roula Said-Nahal<sup>1</sup>, Maria-Antonietta D'Agostino<sup>1,2</sup> & Maxime Breban<sup>1,2,3</sup>

1 – Rheumatology Division, Ambroise Paré Hospital (AP-HP), Boulogne-Billancourt, 92100, France

2 – UFR des Sciences de la Santé, Simone Veil, Versailles Saint Quentin en Yvelines University (UVSQ), Montigny-le-Bretonneux, 78180, France;

3 –INSERM U987, UVSQ, Montigny-le-Bretonneux, 78180, France;

**Background** – Radiographic sacroiliitis is an important outcome in SpA and is mandatory to establish a diagnosis of ankylosing spondylitis (AS). However, few data are available on the factors associated with radiographic sacroiliitis and even less on the predictors of progression from nonradiographic SpA to AS.

**Objectives** – to identify factors associated with radiographic sacroiliitis in a cross-sectional study and predictors of progression to AS in a longitudinal study of SpA patients with positive family history.

**Methods** – The studied cohort consisted of 1,358 patients fulfilling the ASAS criteria for SpA and having at least one first- or second-degree SpA-affected relative, who were recruited by the Groupe Français d'Etude Génétique des Spondylarthrites (GFEGS) and followed longitudinally. Anteroposterior radiographs of the pelvis were examined blindly and independently by 2 qualified examiners (RSN, MB), using the modified New York grading system. Definite radiographic sacroiliitis corresponded to a grade  $\geq 2$  bilateral, or grade  $\geq 3$

unilateral. Association between definite radiographic sacroiliitis at inclusion and demographic and clinical variables was examined using univariate and multivariate regression analysis. Of the 448 cases without definite sacroiliitis at inclusion, 160 patients were followed-up (with new pelvic X-ray) for 2 to 15 years. Factors associated with progression to definite sacroiliitis were assessed by regression analysis in univariate analysis only because of the smaller study sample

**Results** – In multivariate analysis, factors positively associated with definite sacroiliitis at inclusion were male sex ( $p=7 \times 10^{-8}$ ), younger age at disease onset ( $p=0.04$ ), longer disease duration ( $p=0.0005$ ), HLA-B27 positivity ( $p=0.02$ ), inflammatory back pain ( $p=0.02$ ) and uveitis ( $p=9.8 \times 10^{-5}$ ). The only negatively associated factor was enthesitis ( $p=0.002$ ). In longitudinal study, 38 of the 160 patients (23.75%) developed definite sacroiliitis after a mean follow-up duration of  $8.8 \pm 3.0$  years. The only baseline factor associated with progression to definite sacroiliitis throughout follow-up was a low grade radiographic sacroiliitis at inclusion ( $p=0.002$ ). Uveitis and buttock pain occurring during the follow-up period were also associated with progression to definite sacro-iliitis ( $p=0.04$  and  $p=0.05$ , respectively).

**Conclusions** – After an average follow-up duration of almost 9 years, 23.75% of the patients with non-radiographic SpA had developed definitive sacroiliitis (i.e. AS). Factors associated with definitive radiographic sacroiliitis were male sex, younger age at disease, longer disease duration, HLA-B27 positivity, inflammatory back pain, uveitis and lack of enthesitis in the baseline cross-sectional analysis, and low grade radiographic sacroiliitis at inclusion, uveitis and buttock pain during the follow-up in the longitudinal follow-up study.

## **Annexe 3**

---

Article publié dans « La Revue du Rhumatisme – Monographies » en septembre 2014

### **GÉNÉTIQUE : LES NOUVEAUX ÉLÉMENTS**

### **NEW INSIGHTS IN GENETICS**

Félicie Costantino<sup>1,2,3</sup> & Maxime Breban<sup>1,2,3</sup>

- 1- Service de Rhumatologie, Hôpital Ambroise Paré (AP-HP) ; 9 avenue Charles de Gaulle, 92100 Boulogne-Billancourt*
- 2- Institut Cochin, INSERM U1016, CNRS UMR 8104, Université Paris Descartes, Laboratoire d'excellence INFLAMEX ; 27 rue du Faubourg Saint Jacques, 75014 Paris*
- 3- UFR des Sciences de la santé Simone Veil, Université de Versailles-Saint Quentin ; 2 avenue de la source de la Bièvre, 78180 Montigny-le-Bretonneux*

#### Auteur correspondant:

Maxime Breban

Hôpital Ambroise Paré, 9 ave Charles de Gaulle, 92100, Boulogne, France

Ph: 33-(0)149 095 672; Fax: 33-(0)149 095 865

E-mail: [maxime.breban@apr.aphp.fr](mailto:maxime.breban@apr.aphp.fr)

Mots-clés : spondyloarthrite, génétique, HLA-B27, ERAP1, IL23R

Keywords: spondyloarthritis, genetics, HLA-B27, ERAP1, IL23R

## RÉSUMÉ

La spondyloarthrite (SpA) est une maladie multifactorielle complexe possédant une forte composante héréditaire, reflet d'un terrain génétique prédisposant. Les études familiales ont permis d'estimer l'héritabilité de la maladie à plus de 90% et d'établir son caractère multigénique. En effet, si l'antigène HLA-B27 est le facteur majeur de susceptibilité à la maladie, expliquant à lui seul 25 à 50% de la prédisposition génétique, plus de 20 autres locus de susceptibilité à la maladie ont été identifiés ces dernières années grâce à de vastes études d'association pan-génomiques (Genome Wide Association Studies ou GWAS).

Un grand nombre de polymorphismes de susceptibilité sont partagés entre la SpA et d'autres maladies inflammatoires qui lui sont fréquemment associées comme les maladies inflammatoires chroniques de l'intestin ou le psoriasis. Par ailleurs, des études génétiques s'intéressant à l'ensemble des sous-types de SpA ont permis de confirmer certaines des associations déjà mises en évidence dans la spondylarthrite ankylosante (SA), sous-type emblématique de la maladie. L'ensemble de ces éléments est en faveur d'un terrain génétique partagé entre ces différentes pathologies.

Ces découvertes ont également permis d'identifier ou de confirmer de nouvelles voies intervenant dans la survenue de la SpA comme la voie de différenciation lymphocytaire Th17 ou la famille des aminopeptidases. Des études fonctionnelles ont également démontré les conséquences de certains polymorphismes de susceptibilité (comme le polymorphisme rs30187 dans le gène *ERAP1* ou rs11209026 dans le gène *IL23R*) permettant de mieux comprendre l'implication des locus dans la maladie. Cependant, le polymorphisme causal reste à identifier pour la majorité des locus de susceptibilité.

Malgré les découvertes récentes, l'ensemble des variants identifiés à ce jour n'explique qu'une très faible partie de la prédisposition génétique totale et le défi actuel consiste à identifier ce qui constitue encore la part « d'héritabilité manquante ». Le développement rapide du séquençage haut débit pourrait aider à la découverte de variants rares, difficiles à identifier par les approches GWAS classiques.



## ABSTRACT

Spondyloarthritis (SpA) is a chronic complex disease with high heritability reflecting a predisposing genetic background. Heritability has been estimated around 90% and familial studies have established a polygenic transmission of the disease. The most important part of heritability comes from the HLA-B27 allele (explaining 25 to 50% of the heritability). More recently, large genome wide association studies (GWAS) have identified more than 20 other susceptibility loci.

There is a substantial overlap between the loci associated with SpA and other inflammatory diseases frequently associated to SpA, such as inflammatory bowel disease or psoriasis. Moreover recent genetic association studies have demonstrated that some of the polymorphisms associated with AS, the prototypical form of SpA, are also associated with the whole SpA. These results are in favour of a shared genetic background between these diseases.

All these susceptibility loci highlight several pathways potentially involved in the disease pathogenesis such as the Th17 pathway or genes belonging to aminopeptidase family. Functional studies have also demonstrated consequences of several polymorphisms associated with SpA (such as rs30187 in *ERAP1* or rs11209026 in *IL23R*). However causal polymorphism remains to be identified for most of the susceptibility loci.

Despite the recent discoveries, less than 5% of the heritability is now explained by these loci. The current challenge is to identify “missing heritability”. Next-generation sequencing could be very useful to discover rare variants, hard to identify with conventional GWAS approaches.

## **Introduction**

La spondyloarthrite (SpA) est le rhumatisme inflammatoire chronique le plus fréquent en France après la polyarthrite rhumatoïde avec une prévalence estimée à 0,43% [1]. Elle se caractérise par sa forte agrégation familiale témoignant de la présence d'un terrain génétique prédisposant. La survenue d'arthrites réactionnelles dans les suites d'une infection bactérienne suggère également l'implication de facteurs environnementaux. La SpA peut donc être considérée comme une pathologie au déterminisme complexe médié par de multiples facteurs à la fois génétiques et environnementaux pouvant interagir entre les uns avec les autres.

## **Épidémiologie génétique de la SpA**

### ***Caractérisation phénotypique : SpA ou spondylarthrite ankylosante (SA) ?***

La détermination du phénotype à étudier est cruciale pour interpréter la signification et la validité des études génétiques. La plupart des études génétiques dans le domaine des SpA se sont restreintes à la SA, sous-type emblématique de la maladie. L'avantage d'étudier ce sous-type est d'obtenir un groupe de malades considéré comme homogène, car ayant tous une atteinte radiologique suffisamment évoluée pour rendre le diagnostic indiscutable.

Cependant, certains apparentés de patients atteints de SA développent d'autres sous-types de la maladie comme en témoignent l'étude de familles « multiplex » de SpA (au moins deux atteints par famille) menée par le Groupe Français d'Etude Génétique des Spondylarthrites (GFEGS) [2-4]. La principale conclusion de ces études familiales est que les différents sous-types de SpA peuvent être considérés comme des variations phénotypiques d'une seule maladie. L'étude combinée de l'ensemble des sous-types pourrait donc permettre d'avoir une vision plus complète du terrain génétique en rapport avec la maladie.

### ***La SpA, une maladie ayant une forte composante héréditaire***

Les études familiales ont permis de déterminer le poids des facteurs génétiques dans la survenue de la SpA. En effet, l'indice de récurrence chez les apparentés au premier degré, c'est-à-dire le risque d'être malade si l'on a un parent malade, rapporté au risque dans la population générale, a été évalué à 40 pour la SpA prise dans son ensemble [5] et à 80 pour le sous-type SA [6]. L'étude de jumeaux renforce l'implication des facteurs génétiques dans la SpA puisque le taux de concordance pour la SA est plus élevé chez les jumeaux monozygotes que chez les jumeaux dizygotes [7-9]. L'héritabilité a ainsi pu être estimée à 90%, ce qui est très élevé pour une maladie dite « complexe » [6].

Les données familiales ont également permis d'établir un modèle de transmission de la maladie. Il a ainsi été proposé à la fin des années 1990 un modèle impliquant de 3 à 9 gènes (incluant le HLA-B27) exerçant entre eux des effets multiplicatifs [6]. Au regard des découvertes récentes, le nombre total de gènes en cause semble toutefois beaucoup plus important qu'initialement supposé, et le modèle exact de transmission reste à préciser.

## **Facteurs génétiques de susceptibilité à la SpA avant l'ère des Genome Wide Association Studies (GWAS)**

### ***HLA-B27, un facteur prépondérant dans la SpA***

L'association de l'antigène HLA-B27 avec la SA a été découverte au début des années 1970 [10,11] puis confirmée pour la plupart des autres sous-types de SpA [12]. Il s'agit d'une des plus fortes associations entre une molécule HLA et une maladie. En effet le risque de développer une SpA chez les porteurs de l'allèle HLA-B27 est multiplié environ par 40 par rapport aux non-porteurs [1]. L'implication du HLA-B27 dans la survenue de la SpA a été renforcée par le modèle de rat transgénique pour cette molécule qui développe une symptomatologie très proche de celle décrite chez les patients (atteintes inflammatoires articulaires, digestives et cutanées) [13]. Toutefois, le mécanisme de cette association reste encore largement incompris. Nous aborderons cette question dans l'article de cette monographie consacré à l'immunopathologie de la maladie.

HLA-B27 est donc un acteur majeur dans la pathogénie de la SpA. Il n'est cependant ni nécessaire, ni suffisant à sa survenue. En effet, près de 20% des malades ne sont pas porteurs de cet allèle [1]. De plus, parmi les porteurs du HLA-B27, seuls 3 à 5% développent une SpA [1,14]. Ce pourcentage passe à 20% chez les patients HLA-B27+ apparentés à un malade suggérant la présence d'autres facteurs génétiques partagés [14]. L'étude de paires de jumeaux B27+ conforte cette hypothèse puisque la différence du taux de concordance entre jumeaux monozygotes et dizygotes persiste si l'on ne considère que les individus B27+ [7].

### ***Etudes de liaison***

Afin d'identifier des gènes de susceptibilité autres que HLA-B27, les chercheurs ont tout d'abord utilisé les méthodes développées pour les maladies mendéliennes monogéniques parmi lesquelles les analyses de liaison. Ces analyses consistent à étudier la transmission de polymorphismes répartis sur l'ensemble du génome au sein de familles multiplex. Une liaison existe si la distance entre le gène de susceptibilité (et donc indirectement la maladie) et le

polymorphisme étudié est suffisamment faible pour qu'ils soient transmis ensemble d'une génération à l'autre, sans recombinaison.

Trois analyses de liaison pan-génomiques ont été publiées au début des années 2000, deux dans la SA et une dans la SpA [15-17]. Elles n'ont révélé que deux locus significativement liés à la maladie en dehors du complexe majeur d'histocompatibilité (CMH) : 16q et 9q31-34 (**tableau 1**). Cette dernière région ayant été identifiée dans deux des études, elle a été dénommée SPA2 (le locus SPA1 correspondant à la région du complexe majeur d'histocompatibilité comprenant HLA-B27). Une étude approfondie de cette région a permis la découverte d'un haplotype de six SNPs à proximité de *TNFSF15* et d'un SNP intronique dans *TNFSF8* significativement associés à la maladie [18,19].

### ***Etudes d'association gène-candidat***

Les études d'association sont une approche complémentaire aux analyses de liaison. Ce type d'étude consiste à comparer entre des patients indépendants (sans lien de parenté entre eux) et des témoins la fréquence de variants génétiques situés dans des gènes dont la fonction pourrait jouer un rôle dans la SpA. De nombreuses études de ce type ont été réalisées dans la SpA, le choix des gènes testés s'étant basé sur des connaissances physiopathologiques préalables, sur l'étude de modèles animaux ou encore sur le rapprochement fait avec d'autres pathologies proches comme la maladie de Crohn ou le psoriasis. Malgré leur grand nombre, ces études ont été globalement peu fructueuses, se heurtant le plus souvent à la difficulté de confirmer les résultats dans une deuxième population.

## **L'ère des GWAS**

### ***Facteurs de susceptibilité identifiés***

Les études gène-candidat partent d'une hypothèse physiopathologique *a priori* et se limitent généralement à ce qui est déjà connu de la pathologie. Le développement des méthodes de génotypage à haut débit a permis de s'affranchir de cette nécessité et de réaliser des études portant sur l'ensemble du génome. La majorité des gènes de susceptibilité à la SpA en dehors du HLA-B27 ont été identifiés grâce à ces GWAS.

Quatre études de ce genre ont été réalisées dans le sous-type SA, 3 dans des populations caucasiennes [20-22] et une dans une population asiatique (Chinois de l'ethnie

Han) [23]. Ces études ont permis d'identifier 13 locus de susceptibilité à la SA en dehors du CMH chez les Caucasiens et 2 locus additionnels chez les Chinois Han (**tableau 2**).

Plus récemment ont été publiés les résultats d'une étude gènes-candidats haut-débit analysant simultanément un grand nombre de locus potentiellement impliqués dans les maladies auto-immunes ou inflammatoires. Cette vaste étude internationale a inclus plus de 10000 patients atteints de SA et 15000 témoins. Outre la confirmation de la plupart des facteurs génétiques identifiés lors des précédents GWAS, 13 nouveaux locus de susceptibilité ont été identifiés (**tableau 2**) [24].

### ***Que nous apprennent ces nouveaux facteurs sur la SpA?***

Plusieurs points sont à souligner concernant les résultats des GWAS. Tout d'abord, de nombreux gènes de susceptibilité appartiennent à la voie de différenciation lymphocytaire Th17 fortement associée aux réponses pro-inflammatoires (*IL23R*, *IL12B*, *STAT3*, *TYK2*, *IL6R*, *IL27*, *IL1R2/IL1R1*). L'implication de cette voie dans la SpA est renforcée par plusieurs observations in vivo que ce soit dans des modèles animaux [25,26] ou chez l'homme [27]. Plusieurs essais thérapeutiques cherchant à bloquer cette voie sont en cours avec des résultats encourageants notamment pour le secukinumab (anti-IL17) [28].

La forte association de plusieurs polymorphismes d'*ERAPI* à la SpA suggère également l'implication du processus de présentation antigénique de peptides par le CMH de classe I dans la survenue de la maladie. Ce gène code pour une aminopeptidase du réticulum endoplasmique (RE) qui a comme principale fonction de couper les peptides à une longueur optimale pour permettre leur présentation par le CMH de classe I. L'implication du processus de présentation antigénique et de l'apprêtement des peptides est renforcée par l'association à la SA de polymorphismes d'autres aminopeptidases (*ERAP2* et *LNPEP* et *NPEPPS*) [24,29] (**figure 1**). De façon intéressante, une forte interaction génétique a été mise en évidence entre le SNP rs30187 du gène *ERAPI* et HLA-B27 [22,24]. En effet, l'allèle T de ce SNP ne prédispose à la SpA que chez les porteurs du HLA-B27.

Il faut enfin noter que l'ensemble ces nouveaux locus n'explique qu'une très faible fraction de la prédisposition génétique à la maladie : en effet, la fraction du risque attribuable à l'ensemble des facteurs de susceptibilité identifiés en dehors du CMH a été estimée à 4,3% contre 20,1% pour HLA-B27 (**figure 2**) [24]. La majeure partie de la susceptibilité génétique à la SpA reste donc encore à expliquer.

### ***Les facteurs génétiques identifiés dans la SA sont-il des facteurs de susceptibilité à l'ensemble des SpA ?***

Peu d'études génétiques se sont intéressées aux SpA dans leur globalité. Compte-tenu des résultats des études familiales du GFECS et du partage de certains facteurs génétiques avec d'autres pathologies appartenant au spectre des SpA comme la maladie de Crohn ou le psoriasis, notre équipe a cherché à savoir si certains des facteurs génétiques associés à la SA l'étaient aussi aux autres sous-types de SpA [30-32].

Ces études ont montré que certains polymorphismes sont associés à l'ensemble des SpA alors que dans d'autres cas, l'association est restreinte à certains sous-types. Ainsi, les polymorphismes d'*ERAP1* sont associés à l'ensemble des SpA [31]. L'étude de polymorphismes au sein du cluster de gènes de l'*IL1* a quant à elle révélé des différences entre sous-types : le SNP rs1894399 de l'*IL1A* semble plus spécifiquement associé à la SA et le SNP rs3811058 de l'*IL1F10.3* aux formes non ankylosantes de la maladie [30]. Enfin, le SNP non synonyme de l'*IL23R* rs11209026 est fortement associé à la SA alors qu'il n'est pas associé aux formes non ankylosantes de la maladie [32]. L'effet sur l'ankylose semble être en lien avec une modulation de l'inflammation systémique puisque ce SNP est également associé au taux de protéine C-réactive (CRP) dont l'augmentation est un facteur de risque d'ankylose [33]. Les résultats de ces études tendent donc à montrer qu'il existe un socle de facteurs génétiques communs aux SpA alors que d'autres facteurs ont plutôt un rôle de modulateur phénotypique.

### **Le post-GWAS**

#### ***Du polymorphisme à la fonction***

L'un des principaux défis restant à relever est d'identifier les polymorphismes causaux à l'intérieur de chacun des locus de susceptibilité et d'essayer d'en comprendre les conséquences fonctionnelles. L'étude des locus *IL23R* et *ERAP1* a déjà permis des avancées significatives.

Plusieurs polymorphismes du gène du récepteur de l'IL23 ont été associés à la maladie. Le signal d'association semble être expliqué en majeure partie par le SNP rs1120926, dont l'allèle mineur A diminue le risque de développer la maladie [20]. Les études fonctionnelles menées sur ce variant protecteur ont permis de conclure qu'il s'agissait d'un variant perte de fonction entraînant une diminution de la signalisation en aval du récepteur de

l'*IL23* [34-36]. Ces études rejoignent l'observation d'une association entre ce SNP et le niveau de CRP [32] et permettent de proposer un modèle expliquant son association avec l'ankylose (**figure 3**).

Compte-tenu du très fort déséquilibre de liaison dans le gène *ERAP1*, la détermination du polymorphisme causal s'est révélée plus compliquée que pour l'*IL23R*. Toutefois, il semblerait que le signal d'association principal soit porté par le SNP rs30187 avec un signal secondaire attribué à rs17482078 ou rs10050860 [22]. Les études fonctionnelles se sont surtout intéressées à rs30187. Plusieurs études démontrent que l'allèle à risque T entraînerait un gain de fonction avec une augmentation de l'activité enzymatique d'ERAP1 [37,38]. Il reste cependant à déterminer si cet effet ne serait pas lié à un haplotype plutôt qu'à un SNP isolé comme le suggèrent certaines études [31,39,40].

### ***Où trouver l'héritabilité manquante ?***

Malgré le nombre croissant de facteurs de susceptibilité à la SpA découverts ces dernières années, une large part de la prédisposition génétique reste à expliquer et l'un des défis actuels consiste à identifier cette part d'héritabilité «manquante». Le développement rapide du séquençage haut débit pourrait aider à la découverte de variants rares, difficiles à identifier par les approches de GWAS classiques. D'autres approches pourraient également être utiles comme la recherche de variants structuraux (délétions, duplications, inversions, copy number variants), l'analyse des interactions gène/gène ou gène/environnement, ou encore la mise en évidence de variations épigénétiques (méthylation, micro-ARNs,...).

### **Conclusions**

La SpA est donc une pathologie multifactorielle complexe, résultant de l'interaction de facteurs génétiques et environnementaux. Longtemps, l'antigène HLA-B27 a été le seul gène associé de façon certaine à la maladie. S'il reste aujourd'hui le facteur majeur de susceptibilité à la SpA, d'autres facteurs génétiques ont été identifiés, notamment grâce aux avancées technologiques permettant des études sur l'ensemble du génome. Ces découvertes ont permis de mieux comprendre la physiopathologie de la maladie et ouvrent la voie à de nouvelles pistes thérapeutiques. Cependant, de nombreux défis restent à relever. En effet, l'ensemble des facteurs génétiques identifiés à ce jour ne représentent qu'une faible proportion de l'héritabilité totale. De plus, les conséquences fonctionnelles des polymorphismes de susceptibilité sont encore largement méconnues.

## REFERENCES

1. Costantino F, Talpin A, Said-Nahal R, et al. Prevalence of spondyloarthritis in reference to HLA-B27 in the French population: results of the GAZEL cohort. *Ann Rheum Dis* Published Online First: 18 December 2013. doi:10.1136/annrheumdis-2013-204436
2. Said-Nahal R, Miceli-Richard C, Berthelot JM, et al. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. *Groupe Français d'Etude Génétique des Spondylarthropathies. Arthritis Rheum* 2000;43:1356–65. doi:10.1002/1529-0131(200006)43:6<1356::AID-ANR20>3.0.CO;2-Y
3. Said-Nahal R, Miceli-Richard C, D'Agostino MA, et al. Phenotypic diversity is not determined by independent genetic factors in familial spondylarthropathy. *Arthritis Rheum* 2001;45:478–84.
4. Porcher R, Said-Nahal R, D'Agostino M-A, et al. Two major spondylarthropathy phenotypes are distinguished by pattern analysis in multiplex families. *Arthritis Rheum* 2005;53:263–71. doi:10.1002/art.21070
5. Dernis E, Said-Nahal R, D'Agostino M-A, et al. Recurrence of spondylarthropathy among first-degree relatives of patients: a systematic cross-sectional study. *Ann Rheum Dis* 2009;68:502–7. doi:10.1136/ard.2008.089599
6. Brown MA, Laval SH, Brophy S, et al. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann Rheum Dis* 2000;59:883–6.
7. Brown MA, Kennedy LG, MacGregor AJ, et al. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum* 1997;40:1823–8. doi:10.1002/1529-0131(199710)40:10<1823::AID-ART15>3.0.CO;2-1
8. Pedersen OB, Svendsen AJ, Ejstrup L, et al. On the heritability of psoriatic arthritis. Disease concordance among monozygotic and dizygotic twins. *Ann Rheum Dis* 2008;67:1417–21. doi:10.1136/ard.2007.078428
9. Järvinen P. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis Rheum* 1995;38:381–3.
10. Brewerton DA, Hart FD, Nicholls A, et al. Ankylosing spondylitis and HL-A 27. *Lancet* 1973;1:904–7.
11. Caffrey MF, James DC. Human lymphocyte antigen association in ankylosing spondylitis. *Nature* 1973;242:121.
12. Reveille JD. The genetic basis of spondyloarthritis. *Ann Rheum Dis* 2011;70 Suppl 1:i44–50. doi:10.1136/ard.2010.140574



13. Breban M, Hacquard-Bouder C, Falgarone G. Animal models of HLA-B27-associated diseases. *Curr Mol Med* 2004;4:31–40.
14. Van der Linden SM, Valkenburg HA, de Jongh BM, et al. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27:241–9.
15. Laval SH, Timms A, Edwards S, et al. Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. *Am J Hum Genet* 2001;68:918–26. doi:10.1086/319509
16. Zhang G, Luo J, Bruckel J, et al. Genetic studies in familial ankylosing spondylitis susceptibility. *Arthritis Rheum* 2004;50:2246–54. doi:10.1002/art.20308
17. Miceli-Richard C, Zouali H, Said-Nahal R, et al. Significant linkage to spondyloarthropathy on 9q31-34. *Hum Mol Genet* 2004;13:1641–8. doi:10.1093/hmg/ddh179
18. Zinovieva E, Bourgain C, Kadi A, et al. Comprehensive linkage and association analyses identify haplotype, near to the TNFSF15 gene, significantly associated with spondyloarthritis. *PLoS Genet* 2009;5:e1000528. doi:10.1371/journal.pgen.1000528
19. Zinovieva E, Kadi A, Letourneur F, et al. Systematic candidate gene investigations in the SPA2 locus (9q32) show an association between TNFSF8 and susceptibility to spondylarthritis. *Arthritis Rheum* 2011;63:1853–9. doi:10.1002/art.30377
20. Burton PR, Clayton DG, Cardon LR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–37. doi:10.1038/ng.2007.17
21. Reveille JD, Sims A-M, Danoy P, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010;42:123–7. doi:10.1038/ng.513
22. Evans DM, Spencer CCA, Pointon JJ, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–7. doi:10.1038/ng.873
23. Lin Z, Bei J-X, Shen M, et al. A genome-wide association study in Han Chinese identifies new susceptibility loci for ankylosing spondylitis. *Nat Genet* 2012;44:73–7. doi:10.1038/ng.1005
24. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* Published Online First: 9 June 2013. doi:10.1038/ng.2667

25. Sherlock JP, Joyce-Shaikh B, Turner SP, et al. IL-23 induces spondyloarthropathy by acting on ROR- $\gamma$ t<sup>+</sup> CD3<sup>+</sup>CD4<sup>+</sup>CD8<sup>-</sup> enthesal resident T cells. *Nat Med* 2012;18:1069–76. doi:10.1038/nm.2817
26. Glatigny S, Fert I, Blaton MA, et al. Proinflammatory Th17 cells are expanded and induced by dendritic cells in spondylarthritis-prone HLA-B27-transgenic rats. *Arthritis Rheum* 2012;64:110–20. doi:10.1002/art.33321
27. Shen H, Goodall JC, Hill Gaston JS. Frequency and phenotype of peripheral blood Th17 cells in ankylosing spondylitis and rheumatoid arthritis. *Arthritis Rheum* 2009;60:1647–56. doi:10.1002/art.24568
28. Baeten D, Baraliakos X, Braun J, et al. Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. *Lancet* Published Online First: 12 September 2013. doi:10.1016/S0140-6736(13)61134-4
29. Parkes M, Cortes A, Heel DA van, et al. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet* 2013;14:661–73. doi:10.1038/nrg3502
30. Monnet D, Kadi A, Izac B, et al. Association between the IL-1 family gene cluster and spondyloarthritis. *Ann Rheum Dis* 2012;71:885–90. doi:10.1136/annrheumdis-2011-200439
31. Kadi A, Izac B, Said-Nahal R, et al. Investigating the genetic association between ERAP1 and spondyloarthritis. *Ann Rheum Dis* 2013;72:608–13. doi:10.1136/annrheumdis-2012-201783
32. Kadi A, Costantino F, Izac B, et al. The IL23R non-synonymous polymorphism rs11209026 is associated with radiographic sacroiliitis in spondyloarthritis. *Arthritis Rheum* Published Online First: 1 July 2013. doi:10.1002/art.38060
33. Poddubnyy D, Rudwaleit M, Haibel H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. *Ann Rheum Dis* 2011;70:1369–74. doi:10.1136/ard.2010.145995
34. Di Meglio P, Di Cesare A, Laggner U, et al. The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PloS One* 2011;6:e17160. doi:10.1371/journal.pone.0017160
35. Pidasheva S, Trifari S, Phillips A, et al. Functional studies on the IBD susceptibility gene IL23R implicate reduced receptor function in the protective genetic variant R381Q. *PloS One* 2011;6:e25038. doi:10.1371/journal.pone.0025038

36. Sarin R, Wu X, Abraham C. Inflammatory disease protective R381Q IL23 receptor polymorphism results in decreased primary CD4+ and CD8+ human T-cell functional responses. *Proc Natl Acad Sci U S A* 2011;108:9560–5. doi:10.1073/pnas.1017854108
37. Kochan G, Krojer T, Harvey D, et al. Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc Natl Acad Sci U S A* 2011;108:7745–50. doi:10.1073/pnas.1101262108
38. Evnouchidou I, Birtley J, Seregin S, et al. A common single nucleotide polymorphism in endoplasmic reticulum aminopeptidase 2 induces a specificity switch that leads to altered antigen processing. *J Immunol Baltim Md 1950* 2012;189:2383–92. doi:10.4049/jimmunol.1200918
39. Seregin SS, Rastall DPW, Evnouchidou I, et al. Endoplasmic reticulum aminopeptidase-1 alleles associated with increased risk of ankylosing spondylitis reduce HLA-B27 mediated presentation of multiple antigens. *Autoimmunity* Published Online First: 13 September 2013. doi:10.3109/08916934.2013.819855
40. Reeves E, Edwards CJ, Elliott T, et al. Naturally Occurring ERAP1 Haplotypes Encode Functionally Distinct Alleles with Fine Substrate Specificity. *J Immunol Baltim Md 1950* 2013;191:35–43. doi:10.4049/jimmunol.1300598

## TABLEAU ET FIGURES

**Tableau 1** – Résultats des études de liaison pan-génomique réalisées dans la SpA

Etude	Phénotype	Familles	Paires de germains atteints	Locus significatifs	Locus suggestifs
<i>Oxford [15]</i>	SA	185	255	CMH 16q23.3	1p, 2q, 9q, 10q et 19q
<i>NASC [16]</i>	SA	180	244	CMH	Aucun
<i>GFEGS [17]</i>	SpA	120	336	CMH 9q31-34	Aucun

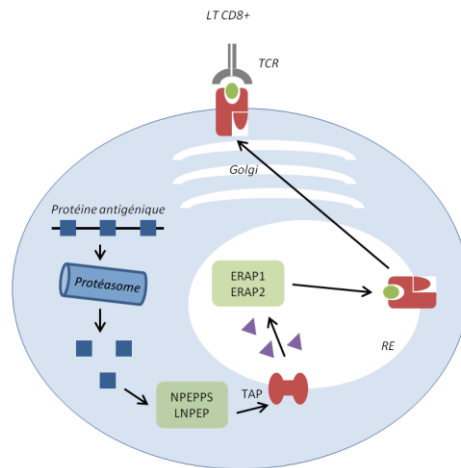
SA : spondylarthrite ankylosante ; SpA : spondylarthrite ; CMH : complexe majeur d'histocompatibilité

**Tableau 2** – Résultats des études d’association pan-génomique réalisées dans la SA

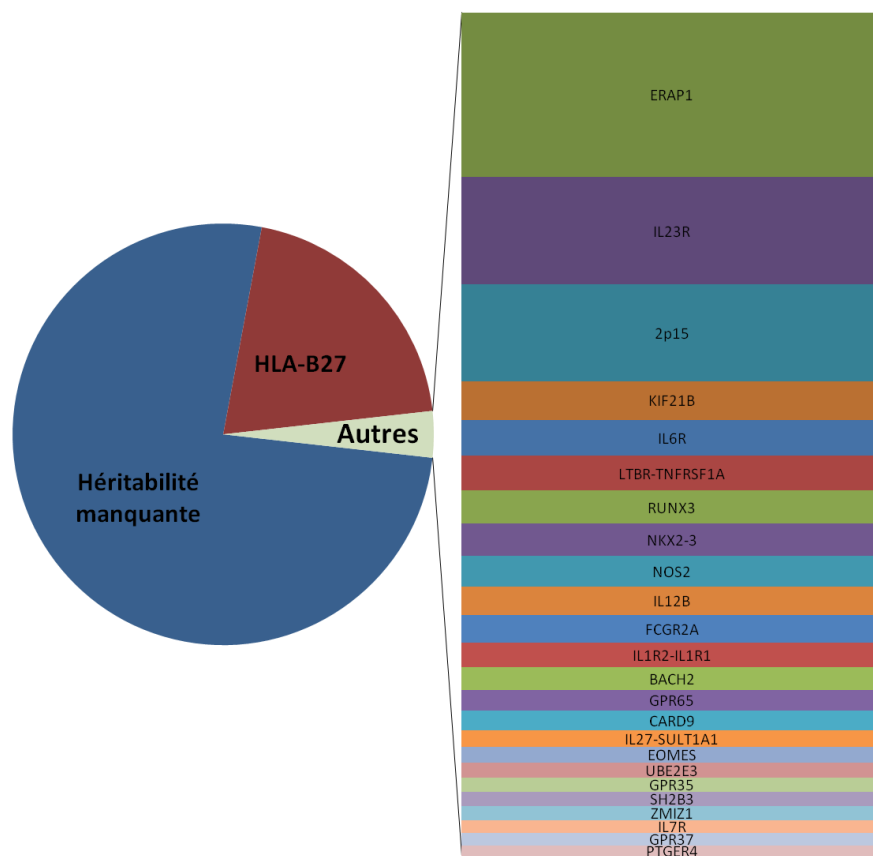
Etude	Population	Nombre de cas/témoins	Facteurs génétiques identifiés (hors CMH)
<i>Burton et al</i> [20]	Caucasienne	Découverte : 922/1466 Réplication : 471/625	<i>IL23R, ERAP1</i>
<i>Reveille et al</i> [21]	Caucasienne	Découverte : 2053/5140 Réplication : 898/1518	<b><i>IL23R, ERAP1</i></b> 2p15, 21q22 <i>ANTXR2, IL1R2</i>
<i>Evans et al</i> [22]	Caucasienne	Découverte : 3023/8779 Réplication : 2111/4483	<b><i>IL23R, ERAP1, 2p15, 21q22,</i></b> <i>KIF21B, RUNX3, IL12B, LTBR-TNFRSF1A,</i> <i>ANTXR2, PTGER4, CARD9, TBKBP1</i>
<i>Lin et al</i> [23]	Chinois Han	Découverte : 1837/4231 Réplication : 2100/3496	<i>EDIL3-HAPLN1, ANO6, 2p15</i>
<i>IGAS et al</i> [24]	Caucasienne et asiatique	Découverte : 10 619/ 15145	<b><i>RUNX3, IL23R, GPR25-KIF21B, 2p15, PTGER4, ERAP1,</i></b> <b><i>IL12B, CARD9, LTBR-TNFRSF1A, NPEPPS, 21q22</i></b> <i>IL6R, FCGR2A, UBE2E3, GPR35, BACH2, ZMIZ1, NKX2-3,</i> <i>SH2B3, GPR65, IL27-SULT1A1, NOS2, TYK2, ICOSLG</i>

Les locus en gras correspondent aux locus déjà associés à la SA dans au moins un précédent GWAS. Les locus en bleu correspondent à des associations suggestives.

SA : spondylarthrite ankylosante ; CMH : Complexe Majeur d’Histocompatibilité



**Figure 1** – Implication des différentes aminopeptidases associées à la SA dans l’apprêtement des peptides avant présentation par le CMH de classe I (d’après Parkes et al [29]) ; RE : réticulum endoplasmique



**Figure 2** – Fraction de l’héritabilité à la SA attribuable à chaque locus de susceptibilité (d’après IGAS et al [24]).



**Figure 3** – Modèle proposé de mécanisme reliant le polymorphisme rs11209026 à la susceptibilité à la sacro-iliite radiologique dans la SpA